EVOLUTIONARY MEDICINE • 100 YEARS OF MAGNETIC MEMORIES • QUANTUM GLUE

SCIENTIFIC AMERICAN

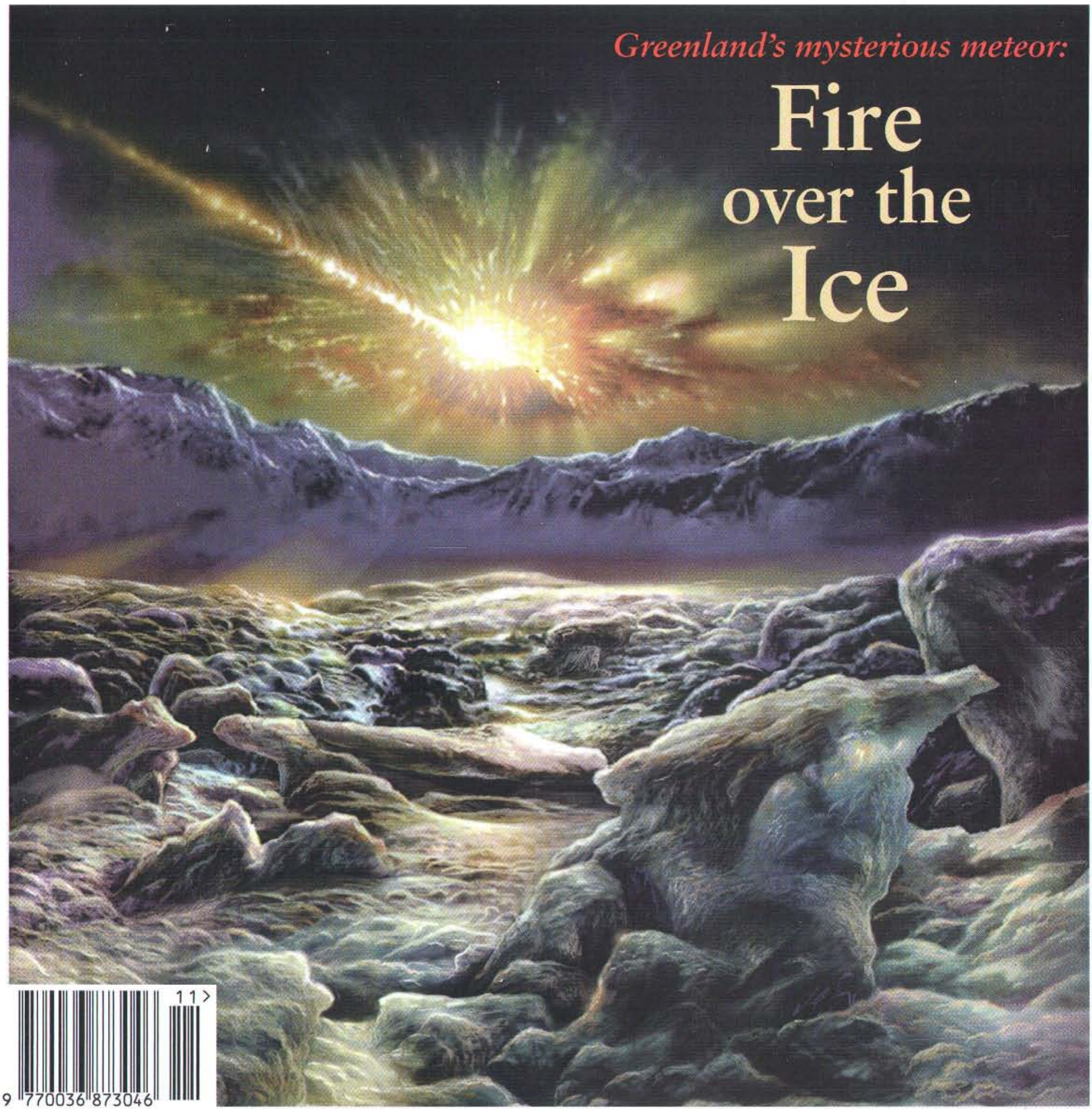
NOVEMBER 1998 $4.95 U.K. £3.20

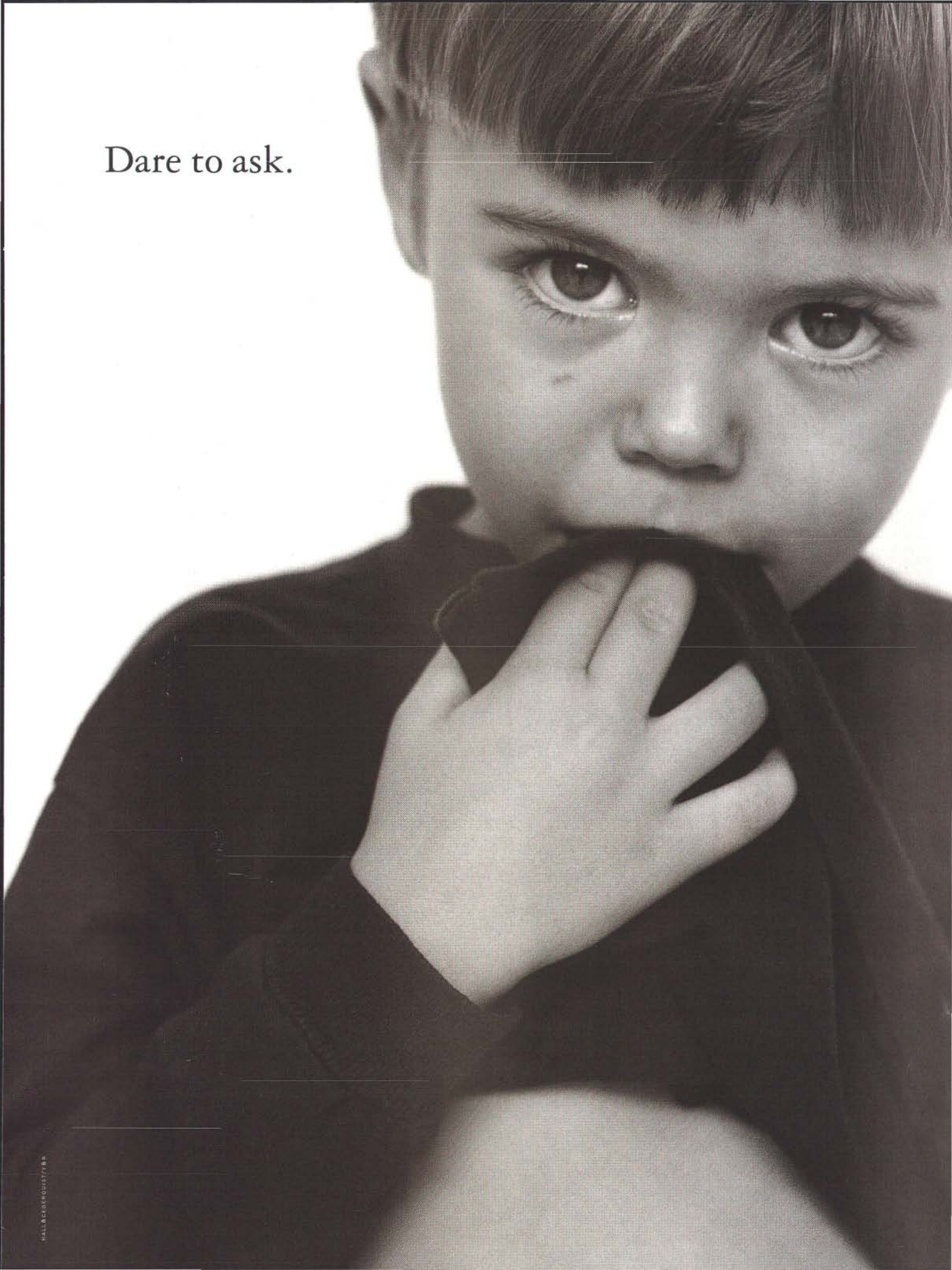
HELL FROM THE HEAVENS

Meteorite impact in the desert turns sand to glass

Greenland's mysterious meteor:

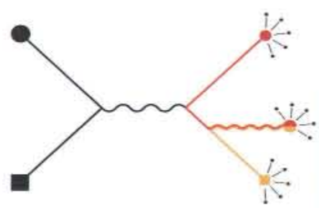
Fire over the Ice

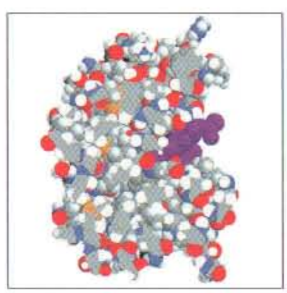


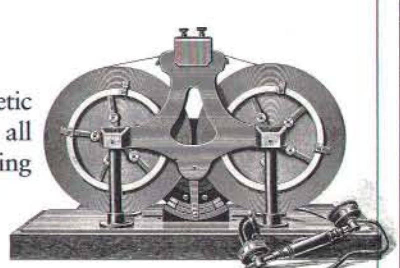

Dare to ask.

SCIENTIFIC AMERICAN

November 1998 Volume 279 Number 5







Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. Periodicals postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 242764. Canadian BN No. 127387652RT; QST No. Q1015332537. Subscription rates: one year \$34.97 (outside U.S. \$49). Institutional price: one year \$39.95 (outside U.S. \$50.95). Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537. Reprints available: write Reprint Department, Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111; fax: (212) 355-0408 or send e-mail to sacust@sciam.com **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 247-7631.**

About the Cover
According to witnesses, the meteor that exploded over Greenland last December was bright enough to turn night into day. Recovering fragments has proved difficult. Painting by Don Dixon.

THE SCIENTIFIC AMERICAN WEB SITE
Read still more about the Greenland meteorite-hunting expedition, including excerpts from one astronomer's diary:
www.sciam.com/explorations/1998/080398meteor/index.html

And check out enhanced versions of this month's other articles and departments, linked to further science resources on the World Wide Web.
www.sciam.com

FROM THE EDITORS

Who and What We Lost

The crash of Swissair Flight 111 on September 2 took the lives of 229 people. Three of them were not strangers to *Scientific American*. Epidemiologist Jonathan M. Mann was co-author of "HIV 1998: The Global Picture," which appeared in our July special report on AIDS and HIV. A founder of the World Health Organization's Global Program on AIDS, he was one of the first to point out the pandemic dimensions of the HIV problem and to link it to social and political conditions. Traveling with him was Mary Lou Clements-Mann of Johns Hopkins University, his immunologist colleague and wife, a researcher leading efforts to test vaccines against the virus.

Pierce Gerety of the United Nations High Commission for Refugees was also going to Geneva that night. His connection to us was personal, not professional; he was related to members of our staff. Gerety brought relief, medical and otherwise, to those dispossessed by wars and other disasters. More than an administrator, he was in the field, rescuing people and property, distributing supplies, negotiating for hostages.

Pierce Gerety

Jonathan M. Mann and Mary Lou Clements-Mann

Science comes to life in laboratories. It matures outside. The Manns and Gerety knew firsthand that dry politics and epidemiology add up as the bodies of the sick, wounded, starving and doomed. When vaccines failed, when therapies failed, when our technologies for maiming outstripped the technologies for healing, the Manns and Gerety witnessed the misery. They persevered anyway. Sometimes readers ask why *Scientific American* publishes articles with a political or social edge. Where's the science? The three of them knew.

March had the world biting its nails that asteroid 1997 XF-11 might pass close enough to the earth in 30 years to collide. (Reanalysis promised a comfortable margin for safety.) Then Hollywood staged a summertime double feature, with *Deep Impact* destroying the world by comet in May and *Armageddon* forcing Bruce Willis to miss his daughter's wedding in July. Call 1998 the Year of the Meteorite.

Researchers are grateful to meteorites for delivering samples from deep space and other worlds, such as the famous Martian rocks recovered from Antarctica. Finding them can be arduous, however. Starting on page 36 are stories of two meteorite-hunting expeditions, one in desert heat, one in glacial cold. The movie rights are available, Mr. Spielberg.

John Rennie

JOHN RENNIE, *Editor in Chief*
editors@sciam.com

AP PHOTO (*Gerety*); AP PHOTO/COURTESY OF JOHNS HOPKINS SCHOOL OF PUBLIC HEALTH (*Manns*)

SCIENTIFIC AMERICAN®
Established 1845

John Rennie, EDITOR IN CHIEF

Board of Editors
Michelle Press, MANAGING EDITOR
Philip M. Yam, NEWS EDITOR
Ricki L. Rusting, SENIOR ASSOCIATE EDITOR
ASSOCIATE EDITORS: Timothy M. Beardsley; David A. Schneider; Gary Stix
W. Wayt Gibbs, SENIOR WRITER
Kristin Leutwyler, ON-LINE EDITOR
EDITORS: Mark Alpert; Carol Ezzell; Alden M. Hayashi; Madhusree Mukerjee; George Musser; Sasha Nemecek; Glenn Zorpette
CONTRIBUTING EDITORS: Marguerite Holloway; Steve Mirsky; Paul Wallich

Art
Edward Bell, ART DIRECTOR
Jana Brenning, SENIOR ASSOCIATE ART DIRECTOR
Johnny Johnson, ASSISTANT ART DIRECTOR
Bryan Christie, ASSISTANT ART DIRECTOR
Dmitry Krasny, ASSISTANT ART DIRECTOR
Bridget Gerety, PHOTOGRAPHY EDITOR
Lisa Burnett, PRODUCTION EDITOR

Copy
Maria-Christina Keller, COPY CHIEF
Molly K. Frances; Daniel C. Schlenoff; Katherine A. Wong; Stephanie J. Arthur; Eugene Raikhel; Myles McDonnell

Administration
Rob Gaines, EDITORIAL ADMINISTRATOR
David Wildermuth

Production
Richard Sasso, ASSOCIATE PUBLISHER/VICE PRESIDENT, PRODUCTION
William Sherman, DIRECTOR, PRODUCTION
Janet Cermak, MANUFACTURING MANAGER
Georgina Franco, PRINT PRODUCTION MANAGER
Silvia Di Placido, PREPRESS AND QUALITY MANAGER
Madelyn Keyes, CUSTOM PUBLISHING MANAGER
Norma Jones, ASSISTANT PROJECT MANAGER
Carl Cherebin, AD TRAFFIC

Circulation
Lorraine Leib Terlecki, ASSOCIATE PUBLISHER/CIRCULATION DIRECTOR
Katherine Robold, CIRCULATION MANAGER
Joanne Guralnick, CIRCULATION PROMOTION MANAGER
Rosa Davis, FULFILLMENT MANAGER

Business Administration
Marie M. Beaumonte, GENERAL MANAGER
Alyson M. Lane, BUSINESS MANAGER
Constance Holmes, MANAGER, ADVERTISING ACCOUNTING AND COORDINATION

Electronic Publishing
Martin O. K. Paul, DIRECTOR

Ancillary Products
Diane McGarvey, DIRECTOR

Chairman and Chief Executive Officer
John J. Hanley

Co-Chairman
Rolf Grisebach

President
Joachim P. Rosler

Vice President
Frances Newburg

SCIENTIFIC AMERICAN, INC.
415 Madison Avenue
New York, NY 10017-1111
(212) 754-0550

PRINTED IN U.S.A.

LETTERS TO THE EDITORS

Readers appreciated the July special report, "Defeating AIDS: What Will It Take?" Dave Toms wrote via e-mail, "Thanks so much for the excellent articles on what's happening with HIV—it's too easy for the comfortable majority (that is, those not directly affected) to forget how bad things still are." And John Casten sent e-mail about taking a copy on a trip to Kathmandu: "I gave it to a friend who works for Family Health International in the HIV/AIDS Prevention and Control Program. He was thrilled to read all the articles with the latest information and passed it around the office." Some readers did have questions, however, about the possibility of developing a vaccine and feasible prevention methods (*below*).

VACCINES AGAINST HIV

In "HIV Vaccines: Prospects and Challenges," by David Baltimore and Carole Heilman, the authors did a commendable job pointing out that both antibodies and cellular immune defenses were most likely important for an HIV vaccine. I disagree, however, with their statement that there is no proof that vaccination against HIV is possible. To the contrary, studies of individuals at very high risk of exposure to HIV have shown that a significant number of these people do not acquire the virus despite multiple, sometimes daily, exposures. This resistance has been correlated with cellular immunity in some cases and, most frequently, with the presence of HIV antibodies in the genital tract. Such antibodies may have preventive value for passive protection against mucosal infections if used, for instance, on a condom or in a spray foam.

JODY BERRY
Department of
Medical Microbiology
University of Manitoba

Baltimore and Heilman reply:

Throughout the article, we tried to identify reasons for optimism that an HIV vaccine may indeed be possible. In particular, we referred to the value of studying individuals who can resist HIV infection despite extensive exposure to the virus. These people, or others who maintain very low levels of HIV infection, may provide information for developing a successful vaccine. Our statement that at present "there is no proof that a vaccine against HIV is possible" was made in the context of that optimism. It referred to our belief that we will only *prove* that the insights we have gained from studying these individuals, as well as from other work, are meaningful when we actually test a vaccine in a large clinical trial. Until then, we can only speculate and hope.

PREVENTION PROGRAMS

Thanks very much for your special report on AIDS. As usual, your coverage provided a thorough and illuminating look at a very important subject. I was particularly impressed that you devoted an entire article to prevention ["Preventing HIV Infection," by Thomas J. Coates and Chris Collins]; however, the bias of the authors against abstinence-based programs left me with several questions. First, the authors state that most people simply will not choose celibacy, yet later they say that sex education caused teens to be less likely to engage in sex. Don't these statements contradict each other? Second, why wasn't the issue of monogamy addressed directly? Finally, the graph on page 77, showing trends in the occurrence of unprotected intercourse, stops with data taken from 10 years ago. Has the trend of improvement continued, leveled off or reversed itself?

DAVID DENNARD
Houston, Tex.

DAN HABIB *Impact Visuals*

SEX EDUCATION PROGRAMS encourage sexually active teens to practice safe sex without causing more teens to have sex.

Coates and Collins reply:

If our article reflects a bias, it is in favor of scientific findings rather than conjecture or hoped-for results. We emphasize the importance of comprehensive sex education because the published, peer-reviewed research indicates that these programs can increase condom use and other self-protective behaviors among young people who choose to have sex and, at the same time, the programs do not lead young people to have increased numbers of sex partners or to initiate sex earlier. There is no such research that abstinence-only programs have positive and sustained effects on the behavior of young people.

Unfortunately, there has been an upward trend in the occurrence of unprotected intercourse during the past two years, probably in part because of the dangerously incorrect thinking that prevention is not necessary once treatments become available. Current medications are far from perfect—prevention is still the order of the day.

A FEW GOOD MEN

In "Where Have All the Boys Gone?" [News and Analysis, July], Mark Alpert reports on a recent paper by Devra Lee Davis about the decline in the male-to-female birth ratio in the U.S. between 1970 and 1990. Davis suggests that the declining ratio is a "sentinel health event" that warns of some environmental hazard. In fact, environmental factors are an unlikely cause. By all measures, the environment is cleaner now than in 1970. Furthermore, the drop between 1970 and 1990 is not unprecedented: the ratio fell faster in the mid-1940s through the late 1950s before rebounding in the 1960s. More pointedly, however, the male-to-female ratio among blacks has actually increased from the mid-1950s to 1994. To suggest that environmental factors are the cause of

changes in the sex ratio would require suppositions about racial differences in the effects of these factors—and that surely runs into Occam's razor.

MICHAEL GOUGH
Cato Institute
Washington, D.C.

MONEY TO BURN

The profile of Stanton A. Glantz by W. Wayt Gibbs ["Big Tobacco's Worst Nightmare," News and Analysis, July] describes Glantz's favoring of a law that "stiffly increases" taxes on cigarettes, reflecting the widely held opinion that such a move would reduce consumption. Although a sudden increase in price or tax on a given item has been shown to reduce its consumption in the short term, it is not at all evident that it does so over the long term.

What Glantz ignores is the well-known phenomenon that expensive items are perceived as prestigious luxury items. Cigarettes in a plain brown wrapper with no logos, no allure and a low price would demonstrate the true value of smoking.

PETER WEBSTER
International Journal of Drug Policy
Le Cannet, France

Letters to the editors should be sent by e-mail to editors@sciam.com or by post to Scientific American, 415 Madison Ave., New York, NY 10017. Letters may be edited for length and clarity.

ERRATA

Because of an editing error, "The Oort Cloud" [September] contains the following incorrect statement: "We have found evidence that a star has passed close to the sun in the past one million years." The sentence should read, "We have found no evidence that a star has passed close to the sun in the past one million years." We regret the confusion.

"Everyday Exposure to Toxic Pollutants" [February] incorrectly indicated that toilet disinfectants are among the major sources of exposure to paradichlorobenzene. The worrisome products containing this chemical are in fact promoted as toilet cleaners or deodorizers, not as disinfectants.

SCIENTIFIC AMERICAN
Advertising and Marketing Contacts

NEW YORK
Kate Dobson, PUBLISHER
tel: 212-451-8522, kdobson@sciam.com
415 Madison Avenue
New York, NY 10017
fax: 212-754-1138

Thomas Potratz, ADVERTISING DIRECTOR
tel: 212-451-8561, tpotratz@sciam.com

Kevin Gentzel
tel: 212-451-8820, kgentzel@sciam.com
Randy James
tel: 212-451-8528, rjames@sciam.com
Stuart M. Keating
tel: 212-451-8525, skeating@sciam.com
Wanda R. Knox
tel: 212-451-8530, wknox@sciam.com

Laura Salant, MARKETING DIRECTOR
tel: 212-451-8590, lsalant@sciam.com
Diane Schube, PROMOTION MANAGER
tel: 212-451-8592, dschube@sciam.com
Susan Spirakis, RESEARCH MANAGER
tel: 212-451-8529, sspirakis@sciam.com
Nancy Mongelli, PROMOTION DESIGN MANAGER
tel: 212-451-8532, nmongelli@sciam.com

ASSISTANTS: May Jung, Beth O'Keeffe

DETROIT
Edward A. Bartley, MIDWEST MANAGER
3000 Town Center, Suite 1435
Southfield, MI 48075
tel: 248-353-4411, fax: 248-353-4360
ebartley@sciam.com
OFFICE MANAGER: Kathy McDonald

CHICAGO
Randy James, CHICAGO REGIONAL MANAGER
tel: 312-236-1090, fax: 312-236-0893
rjames@sciam.com

LOS ANGELES
Lisa K. Carden, WEST COAST MANAGER
1554 South Sepulveda Blvd., Suite 212
Los Angeles, CA 90025
tel: 310-477-9299, fax: 310-477-9179
lcarden@sciam.com
ASSISTANT: Stacy Slossy

SAN FRANCISCO
Debra Silver, SAN FRANCISCO MANAGER
225 Bush Street, Suite 1453
San Francisco, CA 94104
tel: 415-403-9030, fax: 415-403-9033
dsilver@sciam.com
ASSISTANT: Rosemary Nocera

DALLAS
The Griffith Group
16990 Dallas Parkway, Suite 201
Dallas, TX 75248
tel: 972-931-9001, fax: 972-931-9074
lowcpm@onramp.net

International Advertising Contacts

CANADA
Fenn Company, Inc.
2130 King Road, Box 1060
King City, Ontario
L7B 1B1 Canada
tel: 905-833-6200, fax: 905-833-2116
dfenn@canadads.com

EUROPE
Roy Edwards, INTERNATIONAL ADVERTISING DIRECTOR
Thavies Inn House, 3/4, Holborn Circus
London EC1N 2HB, England
tel: +44 171 842-4343, fax: +44 171 583-6221
redwards@sciam.com

BENELUX
Reginald Hoe Europa S.A.
Rue des Confédérés 29
1040 Bruxelles, Belgium
tel: +32-2/735-2150, fax: +32-2/735-7310

MIDDLE EAST
Peter Smith Media & Marketing
Moor Orchard, Payhembury, Honiton
Devon EX14 OJU, England
tel: +44 140 484-1321, fax: +44 140 484-1320

JAPAN
Tsuneo Kai
Nikkei International Ltd.
1-6-6 Uchikanda, Chiyoda-ku
Tokyo 101-0047, Japan
tel: +813-5259-2690, fax: +813-5259-2679

KOREA
Jo, Young Sang
Biscom, Inc.
Kwangwhamun, P.O. Box 1916
Seoul, Korea
tel: +822 739-7840, fax: +822 732-3662

HONG KONG
Stephen Hutton
Hutton Media Limited
Suite 2102, Fook Lee
Commercial Centre Town Place
33 Lockhart Road, Wanchai, Hong Kong
tel: +852 2528 9135, fax: +852 2528 9281

OTHER EDITIONS OF SCIENTIFIC AMERICAN

Le Scienze
Piazza della Repubblica, 8
20121 Milano, ITALY
tel: +39-2-655-4335
redazione@lescienze.it

Spektrum der Wissenschaft
Verlagsgesellschaft mbH
Vangerowstrasse 20
69115 Heidelberg, GERMANY
tel: +49-6221-50460
redaktion@spektrum.com

Investigacion y Ciencia
Prensa Científica, S.A.
Muntaner, 339 pral. 1.ª
08021 Barcelona, SPAIN
tel: +34-93-4143344
precisa@abaforum.es

Pour la Science
Éditions Belin
8, rue Férou
75006 Paris, FRANCE
tel: +33-1-55-42-84-00

Majallat Al-Oloom
Kuwait Foundation for the Advancement of Sciences
P.O. Box 20856
Safat 13069, KUWAIT
tel: +965-2428186

Swiat Nauki
Proszynski i Ska S.A.
ul. Garazowa 7
02-651 Warszawa, POLAND
tel: +48-022-607-76-40
swiatnauki@proszynski.com.pl

Nikkei Science, Inc.
1-9-5 Otemachi, Chiyoda-ku
Tokyo 100-8066, JAPAN
tel: +813-5255-2821

Svit Nauky
Lviv State Medical University
69 Pekarska Street
290010, Lviv, UKRAINE
tel: +380-322-755856
zavadka@meduniv.lviv.ua

Ke Xue
Institute of Scientific and Technical Information of China
P.O. Box 2104
Chongqing, Sichuan
PEOPLE'S REPUBLIC OF CHINA
tel: +86-236-3863170

50, 100 AND 150 YEARS AGO

SCIENTIFIC AMERICAN

NOVEMBER 1948

CYBERNETICS—"Cybernetics is a word invented to define a new field in science. It combines under one heading the study of what in a human context is sometimes loosely described as thinking and in engineering is known as control and communication. In other words, cybernetics attempts to find the common elements in the functioning of automatic machines and of the human nervous system, and to develop a theory which will cover the entire field of control and communication in machines and in living organisms. The word cybernetics is taken from the Greek *kybernetes,* meaning steersman. If the 17th and early 18th centuries were the age of clocks, and the latter 18th and 19th centuries the age of steam engines, the present time is the age of communication and control.—Norbert Wiener"

VIRUS SEX—"Sex was once thought to be the exclusive possession of life's higher forms, yet simpler forms have been found to be possessed of it. Sexual reproduction is the coming together and exchanging of character factors of two parents in making a new individual. Experiments with viruses that attack bacteria showed that inside a bacterium, two or more 'killed' (or mortally damaged) viruses can pool their undamaged parts to make whole individuals capable of reproducing themselves.—Max and Mary Bruce Delbrück"

NOVEMBER 1898

REMOTE CONTROL—"Mr. Nikola Tesla, of New York, has invented what is known in naval science as a dirigible torpedo. Whereas others of the dirigible class use a connecting cable for transmitting controlling power to the torpedo, Mr. Tesla makes use of the Hertzian waves emanating from a distant source (more popularly known as 'wireless telegraphy'), dispensing with the cable. Mr. Tesla is quoted as saying, 'War will cease to be possible when all the world knows that tomorrow the most feeble of the nations can supply itself immediately with a weapon which will render its coast secure and its ports impregnable to the assaults of the united armadas of the world.'"

PROGRESS IN MEDICINE—"We learn from the *Fort Wayne Medical Journal Magazine* for September that at a recent examination before the medical board of Louisiana, Dr. Emma Wakefield, a young negress, passed a successful examination. She is the first woman in the State of Louisiana to study medicine and the first negress in America to receive a medical diploma."

"HOT ZONE" IN VIENNA—"The outbreak of bubonic plague in Vienna due to the experiments in Prof. Nothnagle's bacteriological establishment has spread terror in the Austrian capital. They have several cases in addition to those which resulted in the deaths of Dr. Mueller and Herr Barisch. Extraordinary precautions have now been taken to prevent an epidemic, and everyone who came in contact with Herr Barisch has been isolated. Some of them attempted to escape but were captured and locked up. The plague patients lie in an isolated building and are attended by Dr. Pooch, a volunteer physician, and by Sisters of Charity. It is the opinion of the doctors at the Austrian capital that the plague is likely to spread."

The great telescope at the Observatory of Paris

THE GREAT PARIS TELESCOPE—"The Observatory of Paris is recognized as one of the centers of astronomical work, its astronomers having from the commencement been associated with the history of the science. The great instrument with the staircase shown in our engraving was installed on the grounds in 1875. It is completely inclosed by a metallic cupola (not shown in the engraving). The instrument is provided with a clock movement having a Foucault regulator. The diameter of the mirror is 1.2 meters."

NOVEMBER 1848

A FAMOUS NEUROLOGY CASE—"The Woodstock, Vt., *Mercury* says: 'We gave some account a few weeks ago of the astonishing case of Mr. Gage, foreman of the railroad in Cavendish, who in preparing a charge for blasting a rock had an iron bar driven through his head, entering through his cheek and passing out at the top with a force that carried the bar some yards, after performing its wonderful journey through skull and brains. We refer to this case again to say that the patient not only survives but is much improved. He is likely to have no visible injury but the loss of an eye.'" [*Editors' note: Phineas Gage survived for 12 years but with a radically warped personality; his case is still studied today as a model of cerebral function.*]

NEWS AND ANALYSIS

IN FOCUS

DOGMA OVERTURNED

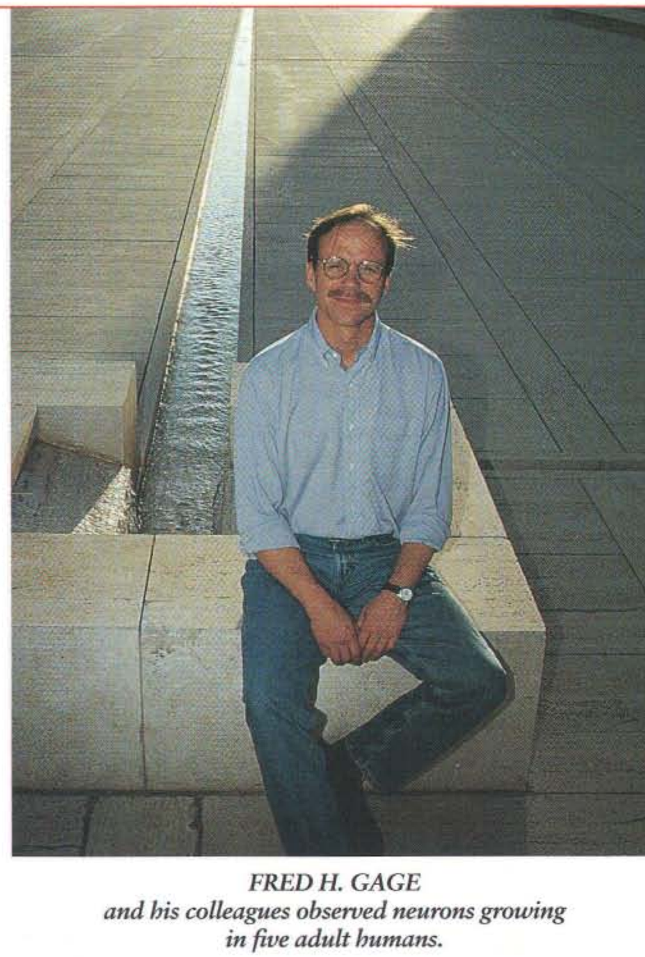
Upending a long-held theory, a study finds that humans can grow new brain neurons throughout life—even into old age

Rats can do it. So can opossums, songbirds, marmosets—why, even tree shrews. But every biology student is taught that humans cannot produce new neurons anywhere in their brains once they have matured. That is a limitation—damage from abuse, disease and injury never heals—but it is also an evolutionary advantage, because it means that memories, imprinted in webs of neurons, can persist undisturbed for a lifetime. Or so the theory has gone for more than a decade.

Now it appears that that fundamental dogma of medicine is wrong; at the very least, it is far too sweeping. Two neuroscientists, one American and one Swedish, have collected the first persuasive evidence that mature, even elderly, people do create additional neurons by the hundreds in at least one important part of the brain, a section of the hippocampus called the dentate gyrus. At press time, the paper was still under review for publication by *Nature Medicine*.

The scientists do not know what the new cells do nor whether the same process, called neurogenesis, occurs elsewhere in the brain. But others in the field say that even though the discovery probably will not yield medical applications for many years, it is a major advance nonetheless. "Once you accept that the brain has some plasticity after all, you have to rethink approaches to lots of problems," says Gerd Kempermann of the University of Regensburg in Germany.

For more than two years, Fred H. Gage of the Salk Institute in San Diego and Peter S. Eriksson of the Göteborg University Institute of Clinical Neuroscience conducted an experiment that was thought to be nearly impossible, for two reasons. First, they needed fresh brain tissue but not from just any spot. The best place to look for newly formed neurons is the hippocampus, which is where they are produced most often in lower mammals. But the hippocampus is nestled



JAMES ARONOVSKY

FRED H. GAGE and his colleagues observed neurons growing in five adult humans.

FIVE AWARDS.

Jean-François Pernette, for exploring some of the world's most inaccessible and spectacular limestone caves.

Amanda Vincent, for conserving seahorse populations in the central Philippines.

Wijaya Godakumbura, for replacing millions of dangerous home-made kerosene lamps in Sri Lanka with safer versions.

Cristina Bubba Zamora, for recovering ancient ceremonial textiles in Bolivia.

Louis Liebenberg, for improving wildlife management in Africa by preserving ancient tracking skills.

deep in the temporal lobes of the brain. "It is very fragile," Eriksson says, and damage to it can destroy a person's ability to learn, because it appears to control which experiences are filed away into long-term storage and which pass into oblivion. Biopsies are thus out of the question.

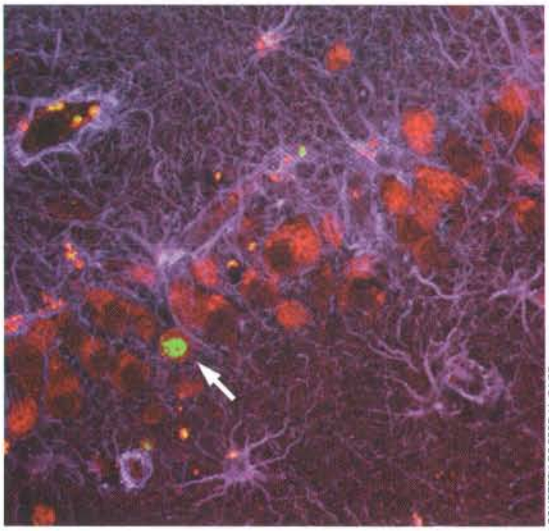
The second problem, Gage explains as he opens a door in his San Diego laboratory to reveal a darkened room full of postdoctoral researchers looking at brain cells through high-tech laser microscopes, is that 60-day-old neurons look just the same as 60-year-old ones. The only well-accepted way to mark nascent cells, neurons or otherwise, is to inject the subject with either tritiated thymidine or bromodeoxyuridine (BrdU), chemicals that can serve as a building block of DNA but that can be detected by film or fluorescence. Cells won't take up these chemicals until they begin to divide and manufacture DNA. When that happens, some of the chemical will be incorporated into the DNA of the offspring, and the young cells will shine for the camera. Unfortunately, both tagging chemicals are toxic to humans. So when Eriksson, on sabbatical at Salk in 1995, began talking to Gage about searching for neurogenesis in humans, there seemed no ethical way to do it.

But after Eriksson returned to Sweden, he found a way. "One day I met this oncologist in the operating room; we were both on call," Eriksson remembers. "I asked him whether he knew anyone giving BrdU to patients, and he said yes; in fact, he knew of a study in which seven people with cancer of the tongue or larynx were getting it." Because newborn cells take up BrdU, researchers can use it to help monitor how fast a tumor is growing.

Eriksson tracked down the doctor in charge of the study, and they made a deal: whenever one of the patients died, the doctor would ask the family's permission to remove the hippocampus. If they agreed, then Eriksson would be summoned. Five times from 1996 until this past February, Eriksson got a call, then jumped in the car and sped over to the hospital to watch as a pathologist pulled out a fingertip-size lump of brain—still warm in one instance—from cadavers aged 57 to 72. He then immediately stained the samples with NeuN, a marker that (as far as is known) attaches only to neurons.

"You need to get the samples within 24 hours, before the cells lose too much of their integrity," Eriksson explains. But the boyish, normally jovial face of the 39-year-old scientist falls as he allows that the work was a touch gruesome. "When your success is based on someone's death, it makes you sad," he says. "It was heartening, though, to tell the families about what good might come from the results of the experiment."

Indeed, the results were surprising. Stepping layer by layer through the stained sections of the dentate gyrus with their laser microscopes, the scientists saw cell after cell lit both green and red. The green meant that the cells had picked up BrdU and thus must have been born while the chemical was in the bloodstream, during the patients' cancer treatments. The red came from NeuN, indicating that the new cells were indeed neurons.



COURTESY OF FRED H. GAGE

FLUORESCENT MARKERS applied to a section of an adult human hippocampus reveal old neurons (red) *and, surprisingly, new ones as well* (green).

"It was an amazing feeling to see them, in every sample, right where we expected they would be," Gage says. "Neurogenesis occurs, and it occurs throughout life. More than that, these new neurons survive for years." One of the patients had received his last BrdU injection 781 days before his death. "Most important," Gage adds, "it is not an isolated, rare event." In all five patients, each cubic millimeter of dentate gyrus held 100 to 300 newly fledged neurons.

That may not sound like a lot, especially considering that the dentate gyrus is no bigger than a BB. But a few neurons can go a long way, Kempermann points out. "Fewer than 50 cells are thought to control breathing," he says; damage to a couple thousand neurons by Parkinson's disease can cause terrible debility.

By the same token, adding a few new neurons to a damaged part of the brain might help the organ repair itself. "That is the real significance of this work," says Pasko Rakic, head of the neurobiology department at Yale University and a chief proponent of the no-new-neurons theory. "To be useful, new neurons must develop connections with their neighbors. [Gage and Eriksson] haven't shown that that happens. And new cells have not been shown in the cerebellum, the cerebral cortex or the thalamus," regions most often damaged by injury or disease. "But this work does suggest the possibility of finding a factor that can encourage cell proliferation elsewhere in the brain."

"It allows us to think about growing neurons for transplantation," Eriksson elaborates. "In experiments at the University of Lund, transplanted fetal cells greatly reduced the symptoms of Parkinson's disease, an effect that lasted for years. But there are ethical concerns with using cells from aborted fetuses." Now there can be reasonable hope of eventually using adult tissue instead.

Such clinical benefits, Eriksson predicts, "are 10 years away, at best." Gage concurs: "Nothing here can be immediately translated to help a person in a wheelchair." That will have to wait until scientists learn much more about where the progenitor cells that give birth to new neurons exist in the brain, what chemical signals spur them to divide, and what determines whether newly created cells become neurons or some other kind of brain matter. Both scientists have animal experiments under way to tackle those tough questions. But it may be years before their peers elsewhere can arrange to get the human tissue needed to confirm their discovery and to build sound medicine on it. So, most likely, "the general spirit of the dogma will live on," Eriksson concedes. "This represents one exception to it; that's all." But where there once seemed only an impenetrable wall, the outline of a door has appeared.

—*W. Wayt Gibbs in San Diego and Göteborg, Sweden*

SCIENCE AND THE CITIZEN

PHYSICS

INCONSTANT CONSTANTS

Do distant galaxies play by different laws of physics?

Of all the assumptions that undergird modern science, perhaps the most fundamental is the uniformity of nature. Although the universe is infinitely diverse, its basic workings appear to be the same everywhere. Otherwise, how could we ever hope to make sense of it? Historically, scientists presupposed uniformity on religious grounds. In this century, Albert Einstein encapsulated it in his principle of relativity. As geologists and astronomers peered far beyond the domain of common experience, they saw no sign that nature behaved any differently in the distant past or in deep space.

Until now. A team of astronomers led by John K. Webb of the University of New South Wales has found the first hint that the laws of physics were slightly different billions of years ago. "The evidence is a little flimsy," says Robert J. Scherrer of Ohio State University. "But if it's confirmed, it'll be the most startling discovery of the past 50 years."

The work is the latest in an effort that began with the musings of English physicist Paul A. M. Dirac in the 1930s. He and others asked whether the constants that appear in their equations—the speed of light in a vacuum, the charge on the electron and so on—are actually constant. Even if the equations themselves are fixed, if the constants varied, nature would have worked in different ways at different times.

But looking for inconstancy is tricky. If the speed of light, for example, were slowly shrinking, we might never know it, because our measuring apparatus might be shrinking, too. For this reason, physicists focus on constants whose values are independent of the measurement system—particularly the fine-structure constant, the ratio of electromagnetic energy to the energy inherent in mass. If it once varied from its present value (roughly $^{1}/_{137}$), subatomic particles would have interacted differently with one another and with light.

Our very existence indicates that the constant is constant or nearly so. If it had varied by more than a factor of 10, carbon atoms would not be stable, and organic life could not have arisen. A tighter case for constancy emerged in the 1970s, when French scientists unearthed the telltale signs of a naturally occurring nuclear reaction in the Oklo uranium deposit in southeastern Gabon. Based on the composition of the nuclear waste, Russian physicist Alexander I. Shlyakhter and others concluded that the constant at the time of the reaction, two billion years ago, was identical to its present value (within the experimental precision of a few parts in 10 million).

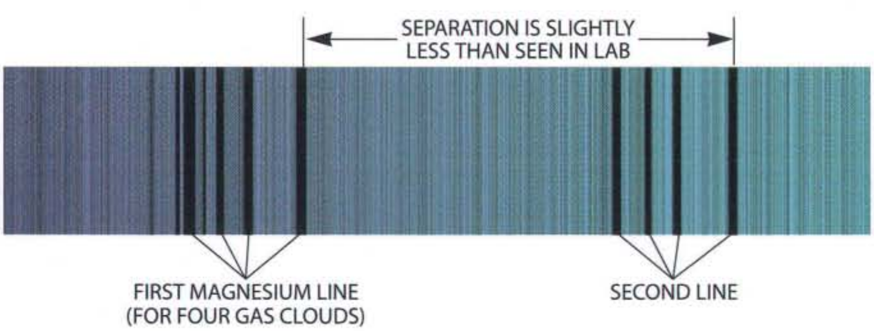
The new finding by Webb's team involves another approach: looking for changes in how atoms absorb light from quasars. These cosmic lighthouses are thought to be powered by massive black holes in far-off galaxies. Their spectra are filled with a forest of thin, black lines, etched when intergalactic gas clouds blocked light of specific wavelengths. If the fine-structure constant has varied, these wavelengths should differ from those measured in the lab.

Although the technique was devised and first applied decades ago, Webb and his colleagues—Victor V. Flambaum and Michael J. Drinkwater, as well as Christopher W. Churchill of Pennsylvania State University and John D. Barrow of the University of Sussex—improved its precision 1,000-fold by simultaneously analyzing spectral lines caused by several chemical elements. They saw no variation in the fine-structure constant over the past seven billion years, which agrees with the Oklo finding. But for the more distant (and hence older) gas clouds, the constant was two parts in 100,000 smaller than today. No known experimental error could mimic the effect.

Still, theorists are skeptical. According to a new paper by Mario Livio and Massimo Stiavelli of the Space Telescope Science Institute in Baltimore, relativity might countenance a slight shift in the fine-structure constant because of cosmic expansion, which dilutes electric charge and therefore reduces electromagnetic energy. Yet any change should be smaller and less abrupt than the observed variation. "It could be that Einstein's equations are wrong, but that is not something you give up lightly," Livio says. "There are very few people, if any, who believe the Webb result."

Even post-Einsteinian physics is stymied. In string theory the fine-structure constant is not fixed; it represents the size of an extra spatial dimension, which we perceive as electromagnetism rather than as another direction. If the dimension somehow changed in size, so would the fine-structure constant, as Thibault Damour of the Institute for Advanced Scientific Study in Bures sur Yvette, France, and Alexander M. Polyakov of Princeton University argued four years ago. But even this effect should be a ten-thousandth of that observed by Webb's team. Other speculations call for the sudden decay of some (unknown) kind of dark matter and the shenanigans of (undetected) gossamer particles.

With the stakes this high, observers are pressing other searches for inconstancy. Scherrer, his graduate student Manoj Kaplinghat and Michael S. Turner of the University of Chicago argue that changes in the constant should show up in the cosmic microwave background radiation. The Microwave Anisotropy Probe satellite, scheduled for launch in



SPECTRAL LINES FROM QUASAR LIGHT *are slightly closer together than the equivalent lines in lab measurements—suggesting that the laws of physics have changed over time.*

CHRISTOPHER W. CHURCHILL *Pennsylvania State University* AND STEVEN S. VOGT *University of California, Santa Cruz*; GEORGE MUSSER AND JOHNNY JOHNSON

2000, should be able to see any variation greater than 1 percent. Though less sensitive than the quasar spectra, this technique probes a much earlier period in cosmic history.

Damour and Lute Maleki and John D. Prestage of the Jet Propulsion Laboratory in Pasadena, Calif., and their colleagues have proposed mounting atomic clocks on a spacecraft and sending it toward the sun. Any variation of the fine-structure constant would shift the frequency of radiation emitted by the atoms on which these clocks are based. Clocks based on different elements would keep different times; the sun's gravity would amplify any discrepancy, making the measurement the most precise ever.

If confirmed, would Webb's findings eventually be explained by a deeper theory, vindicating physicists' faith in a uniform nature? Or would they mean that we live in a frighteningly arbitrary and variegated cosmos, where huge swathes of space abide by alien principles? As physics comes close to unifying its theories, such discrepant observations are becoming less likely—and potentially more momentous. —*George Musser*

PUBLIC HEALTH

THE *E. COLI* ARE COMING

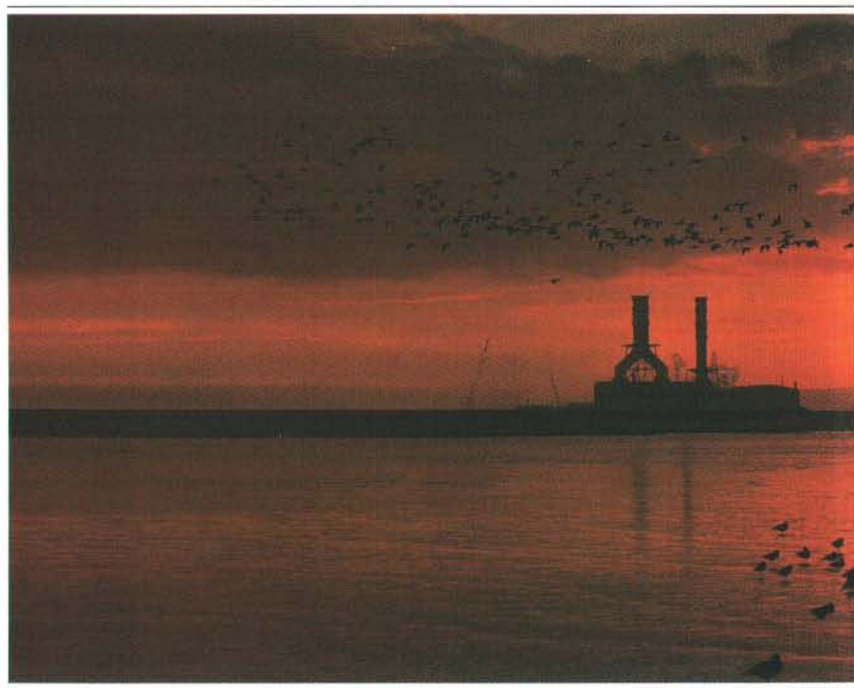
Do toys and toothpaste breed resistant bacteria?

Social critics sometimes proclaim that microbes seem to have supplanted the Soviets as a dire threat to the American way of life. Headlines trumpet the perils of flesh-eating bacteria and the deadly bugs lurking in raw hamburger. Although the U.S. won the cold war, some new evidence suggests that unless we learn to live with them, the bugs may win the battle in the hot zone.

For the past couple of years, Stuart B. Levy of Tufts University has warned of the fallout from undue preoccupation with germ fighting. The proliferation of household products that kill or inhibit bacteria might be helping to create a generation of superbugs that can withstand the chemical onslaught from disinfectants and, in some cases, exhibit resistance to antibiotic drugs.

Now Levy and his colleagues Laura M. McMurry and Margret Oethinger have presented some evidence to stoke those fears. In early August the team published a study in *Nature* that chronicled how a single gene mutation in a laboratory strain of *Escherichia coli* could create a bacterium resistant to triclosan, an antibacterial product used in consumer goods ranging from toothpaste to children's toys. Antibacterials—or more precisely "biocides" because they also kill microorganisms other than bacteria—generally disable a cell in multiple ways; they might, for instance, make a cell membrane permeable while also interfering with enzymes and nucleic acids. This multiple attack makes development of resistance more difficult. But the new Tufts work shows that triclosan may act like an antibiotic, which interferes with only a single cellular process—in particular, it impedes the action of an enzyme essential to synthesizing lipids in the cell wall.

If resistance did develop, it could render ineffective the antiseptic soaps used



in hospitals and in the homes of immunocompromised patients. To be sure, Levy's was a laboratory experiment. And, as could be predicted, the findings evoked an immediate drubbing from industry trade groups, which note that after 30 years of routine triclosan use no evidence of resistance has emerged. A. Denver Russell of Cardiff University in Wales, who has conducted research on antibacterials, some studies of which received industry backing, asserts that it is too soon to draw any conclusions from the Tufts experiment. Triclosan's effects, he says, might result from more than just inhibition of the enzyme, which would make resistance more unlikely. "They haven't looked at other possible mechanisms," Russell says.

Yet Levy continues to amass additional proof of resistance to antibacterials. One journal, *FEMS Microbiology Letters,* has accepted a paper from his laboratory that shows that a number of strains of mutant *E. coli* can overproduce tiny cellular "efflux" pumps that eject chemicals from the cell. This mechanism produces a relatively low level of resistance but is in some ways more insidious. The efflux pumps not only rid the bacterial cells of triclosan but also can flush out antibiotic drugs.

For healthy individuals, Levy says, these chemicals are generally no more useful than ordinary household cleansers. "People are trying to sterilize their environment from bacteria," he says. "But that is only possible in a laboratory. One is only going to remove some bacteria and replace them with others that are insensitive to the antibacterial product and may be potentially harmful." Thus, peaceful coexistence is the only defensible strategy. *—Gary Stix*

MEDICAL DIAGNOSTICS

DOWN DETECTION

New blood and ultrasound tests for Down syndrome might reduce the need for amniocentesis

A drugstore urine test indicates to a 38-year-old woman that she is pregnant. After examining her and taking her history, her gynecologist tells her that she is roughly 10 weeks into the pregnancy. Although the woman is elated, she is also worried about Down syndrome, a form of mental retardation caused by an extra copy of chromosome 21 that occurs more often in the offspring of women older than 35. She and her husband have decided that they would opt for abortion if they conceive a fetus with the disorder. Her doctor says blood tests can determine whether the fetus has Down syndrome but only between weeks 16 and 18 of gestation—during the second trimester. That means the woman might face an abortion in the fifth month, which is particularly traumatic because such late abortions usually involve inducing labor and delivering the fetus.

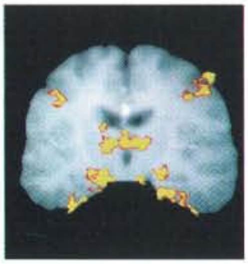
The above scenario occurs hundreds of thousands of times every year in the U.S. alone. But researchers are now evaluating whether a suite of blood tests—one of them new—can be combined with a novel ultrasound technique to detect Down syndrome reliably in fetuses as early as 10 weeks after conception.

The current blood diagnostic for Down is the triple test, which measures the levels of three proteins. One is beta-

IN BRIEF

Unforgettable?

Using advanced functional magnetic resonance imaging (fMRI), two groups of scientists have captured the first images of memories being formed within the brain. Randy L. Buckner of Washington University and his colleagues at Harvard University and Massachusetts General Hospital measured brain activation in young adults as they completed verbal tasks. Later the subjects were asked which words they remembered. James B. Brewer and his colleagues at Stanford University conducted a similar investigation, asking subjects to recall photographs. In both studies, higher levels of activity in the prefrontal and parahippocampal cortices—regions long thought to be involved in encoding memory—corresponded with stronger memories.



REPRINTED FROM SCIENCE (Vol. 281, Aug. 21, 1998)

Winding the Master Clock

The big wheel keeps on turning all right, but not at the same speed. The earth's rotation is actually slowing down. Thus, on December 31 the U.S. Naval Observatory, working for the International Earth Rotation Service (IERS), will add a leap second to the Coordinated Universal Time, the basis for world timekeeping. It is the 22nd leap second added since 1972, when the IERS decided to let atomic clocks—accurate to within a billionth of a second a day—run independently of the earth, which as a clock is only good to about one thousandth of a second a day.

Children's Pollution

Children are often more susceptible to chronic coughing, bronchitis and asthma—and researchers from the University of North Carolina at Chapel Hill think they know why. In a recent study William D. Bennett and his colleagues had children inhale harmless carnauba wax particles and measured the quantity of wax left per unit of lung surface area using a laser device. They found that children retain 35 percent more of the airborne particles they inhale on the surface of their lungs than adolescents or adults do. Similarly sized particles are extremely prevalent in urban air pollution.

More "In Brief" on page 18

RONALD J. WAPNER *Jefferson Medical College*

DOWN SYNDROME FETUS, 12 weeks old, has thicker neck membranes (indicated by arrow) *than those of a healthy baby.*

hCG, part of the human chorionic gonadotropin hormone, which is the protein detected by most pregnancy tests. Beta-hCG is elevated in the blood of women carrying a Down fetus. The other two proteins are estriol, which is lowered in women with a Down pregnancy, and alpha fetoprotein, which is reduced in Down (and elevated in cases of neural-tube defects such as spina bifida).

The triple test detects 60 to 70 percent of Down fetuses, and women who test positive are advised to have amniocentesis to confirm the blood-test result. To perform an amniocentesis, physicians use a long needle to collect fetal cells floating in the amniotic fluid. They then break the cells open and look for any extra chromosomes. But amniocentesis has its risks, including infection, leakage of amniotic fluid and—rarely—the development of clubfoot in the newborn. (Clubfoot is thought to result because amniotic fluid leakage reduces the space the fetus has to develop.)

Laird G. Jackson of Jefferson Medical College in Philadelphia hopes to develop a reliable method for detecting Down syndrome early in pregnancy without the risks of amniocentesis. In preliminary studies, Jackson and his colleagues found that a combination of blood tests and ultrasound detected 90 percent of fetuses with Down syndrome as early as 10 weeks. The blood tests measured beta-hCG and pregnancy-associated plasma protein A (PAPP-A), which scientists have only recently found is decreased in women carrying fetuses with Down. The researchers then used sonography to measure the thickness of the back of the neck of each developing fetus. In 1994 Kypros H. Nicolaides of King's College Hospital in London and his associates had found that an increase in the thickness of the nuchal (neck) membranes suggested Down syndrome.

Jackson says a new, larger study—which will involve 6,000 pregnant women at 12 health centers in the U.S.—should yield results by the end of 2000. But even if the new study confirms the early results, it might not change how most pregnant women are screened for Down syndrome, in part because accurately measuring the nuchal membranes requires precise sonographic techniques that are difficult to standardize.

"There's no question that [nuchal membrane thickness testing] can be useful when it is properly measured," says James E. Haddow of the Foundation for Blood Research in Scarborough, Me. "The problem is, that's hard to do." Earlier this year Haddow and his colleagues published a study in the *New England Journal of Medicine* showing the efficacy of beta-hCG and PAPP-A in diagnosing Down syndrome at 16 health care centers. They could not, however, obtain consistent results using nuchal membrane testing.

Jackson says the new study will use a newly developed, standardized procedure to measure embryonic nuchal membrane thickness. "It's pretty simple and straightforward," he says. "Any skilled radiographer should be able to do it." If all goes well, Jackson hopes that sonogram technicians in obstetric offices across the country will one day be using the technique. —*Carol Ezzell*

FIELD NOTES

TOOL TIME ON CACTUS HILL

In search of the earliest Americans

On a scorching Saturday in late August in southern Virginia, at the end of a dirt track leading through fields of corn and soybeans, archaeologist Michael F. Johnson sits in the shade of oak and hickory trees eating his packed lunch. Nearby, bright-blue tarpaulins protect excavations that have brought Johnson here most weekends for the past several years.

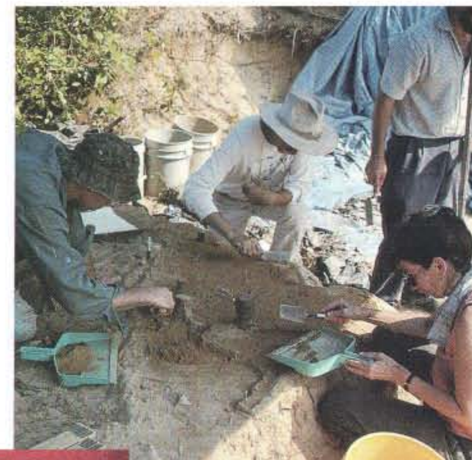
The object of Johnson's passion is a dune of blown sand known as Cactus Hill. Between bites, Johnson is debating with visiting archaeologist Stuart J. Fiedel what the place was like 14,000 years ago. It must have been ideal for a summer camp, Johnson thinks. Facing north,

it would have been cooled by winds coming off glaciers hundreds of miles distant. He offers me an inverted plastic bucket to sit on. The dune would have been dry, he continues, a welcome relief from the surrounding insect-infested bogs. The Nottoway River was at the time only a stone's throw away. There were lots of animals: mastodon, elk, bison, deer, perhaps moose and caribou.

And there were people, maintains Johnson (who is employed by the Fairfax County Park Authority), hunter-gatherers whose descendants may have given rise to Native American tribes. Johnson has found at Cactus Hill quartzite blades, blade fragments and both halves of a broken "point" suitable for a spear, fully nine inches below the well-defined Clovis horizon at the site. That level, recognized all over the country by its characteristic and abundant stone-tool technology, was created 13,000 years ago, according to Fiedel, who conducts surveys for John Milner Associates. (Several studies in the past few years indicate that the conventional date of 11,000 years, based on radiocarbon dating, is a significant underestimate.) Only in recent years has a long investigation at Monte Verde in Chile finally convinced most archaeologists that humans were in the Americas well before Clovis times, so a new potential pre-Clovis site is an important rarity.

In a separate, adjacent dig at Cactus Hill, Joseph M. McAvoy and Lynn D. McAvoy of the Nottoway River Survey have found numerous blade-type tools, some associated with charcoal fragments that tested at 15,000 and 16,000 years old by radiocarbon dating or 18,000 to 19,000 years old by Fiedel's recalibration. Johnson is excited that McAvoy's larger excavation and his own have found "fully comparable" artifacts from below the Clovis horizon. Cactus Hill is "one of the best candidate pre-Clovis sites to come down in a long time," says C. Vance Haynes, Jr., of the University of Arizona, a leading scholar of Paleo-Indian cultures.

On this day Fiedel is listening hard to Johnson's arguments in favor of pre-Clovis occupation, but he is frowning. Johnson says 14,000 years is a "conservative" estimate of the age of his oldest finds. Fiedel agrees that Johnson's fragments are clearly human artifacts, but he is not persuaded by his dates. "You



TIM BEARDSLEY

TIM BEARDSLEY

SEARCHING FOR CLUES to human occupation of Cactus Hill are archaeologist Michael F. Johnson (center, above) *and his helpers. Artifacts found there may be 14,000 years old* (left).

***In Brief**, continued from page 16*

Homeless Orangutans

Forest fires in Borneo have left hundreds of orangutans stranded and starving. Between 1980 and 1990, habitat loss and poaching destroyed up to half of the world population of wild orangutans. Now the fires have claimed another 625,000 acres of the animals' remaining homeland in the virgin forests of East Kalimantan. The Wanariset Orangutan Rescue and Rehabilitation Center, funded by the World Society for the Protection of Animals, currently shelters about 200 orangutans rescued from the fires; the center planned to reintroduce some 30 animals back into intact areas of the forest in September.

MICKEY GIBSON *Animals Animals*

Sticky Soy

This isn't the stuff that ends up on your sleeve after a sushi-eating spree. Kansas State University researcher Xiuzhi Susan Sun has used soy protein to create a water-resistant, nontoxic, formaldehyde-free glue. Sun found a group of nontoxic chemicals that unfold the soy protein molecule, increasing the contact area and thus the molecule's adhesive strength. So far the substance has done well in standard testing: it stayed strong after eight weeks in a chamber at 90 percent relative humidity, proving its value for indoor uses. And as for its ability to withstand the weather, it readily held plywood together after three cycles of being soaked in water for two days and then dried.

Joys of Parenting

Although a recent controversial book claims that parents don't have a big impact on how their kids turn out, other work indicates that parenting does largely influence how the parents turn out. Indeed, John Allman of the California Institute of Technology and his colleagues found that among 10 different primate species, the parent that cared for the offspring significantly outlived the one that didn't—regardless of gender. For instance, male titi monkeys, which care for their children once the female has given birth, outlive their mates by 20 percent.

More "In Brief" on page 20

can't be sure stuff hasn't moved around," he says later. Burrowing wasps and rodents, notoriously, can move objects through sand. McAvoy's published evidence of a pre-Clovis technology at Cactus Hill is "fairly convincing," Fiedel says, but the radiocarbon dates seem almost too old, suggesting evidence of fire 5,000 years before the Clovis culture exploded—a time when few other signs of humans have been documented. Haynes, too, notes that there could be unrecognized errors in the dating of the Cactus Hill layers.

Johnson is undeterred. The pieces of his prized ancient broken point came from the same level but were found several feet apart: because animals would hardly move the separated fragments vertically the same distance, they are probably in their original bed, he argues. Moreover, the stone and the style of workmanship differs from that of Clovis material. "I'm really confident it doesn't fit into Clovis," he says. Johnson's opinion on tool styles counts for something; he has taught himself how to make "Clovis" points that can fool most people.

Fiedel and the other visitors at Cactus Hill this day continue to spin scenarios about the earliest Americans as they take up tools and patiently skim successive half-inch layers of sand from a more recent horizon. Perhaps the inhabitants were members of a hypothetical proto-Clovis culture, Fiedel muses. He observes that some blades like Johnson's and the McAvoys' have recently come to light in South Carolina. But when did the makers arrive from Eurasia? The land bridge that connected it to Alaska was often covered by glaciers. The Cactus Hill archaeologists visiting Johnson's dig, all donating their time, ponder the conundrums as they patiently mark every visible fragment of stone and photograph each exposed level, then sift through the removed material for anything they might have missed the first time. The heat is daunting. As the afternoon wears on, the debate between Johnson and Fiedel moves first one way, then the other, like a tug-of-war.

The debate might never be resolved. The site's owner, Union Camp Corporation, has halted sand mining at Cactus Hill, provided some security and allowed the archaeologists complete access, but time presses. Johnson grimaces as he lifts a tarp to show a ruined trench where the Clovis horizon has been crudely dug out by looters in search of stone points, which can sell for thousands of dollars each. In the process, the pillagers have destroyed layers above and below Clovis. Cactus Hill may be among the earliest inhabited sites in the U.S. But if point rustlers continue to run ahead of the volunteers, science may forever be unable to prove it.

—Tim Beardsley near Petersburg, Va.

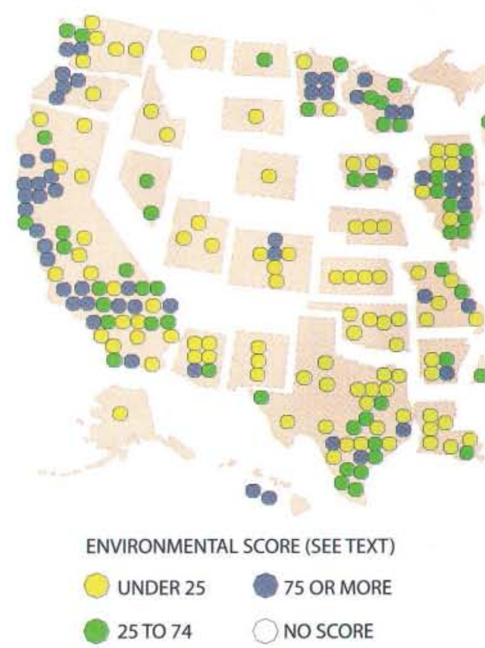
BY THE NUMBERS

How Congress Voted on the Environment

One of the enduring anomalies of political life is that Americans overwhelmingly picture themselves as environmentalists while Congress often votes against environmentalist positions. The explanation for this discrepancy probably lies in several factors: public indifference to arcane but important legislative details; the influence of commercial interests, which make substantial campaign contributions to all parties; and principled opposition by conservative legislators to mandates from Washington.

To give the public more information on environmental voting, the League of Conservation Voters (LCV) was founded in 1970. The maps are based on LCV ratings of members of the 105th Congress through mid-September. In compiling the ratings, the LCV used a select list of 14 environmental measures con-

HOUSE OF REPRESENTATIVES



sidered by the Senate and 29 considered by the House. (Included in the latter is one declaration of co-sponsorship, in which members do not actually vote but express their viewpoint.) The measures chosen represent the consensus of more than two dozen environmental and health groups. These measures, in the words of the LCV, "presented Members of Congress with a real choice on protecting the environment and help distinguish which legislators are working for environmental protection." The LCV scores range from 0 to 100, where high scores represent pro-environmentalist positions.

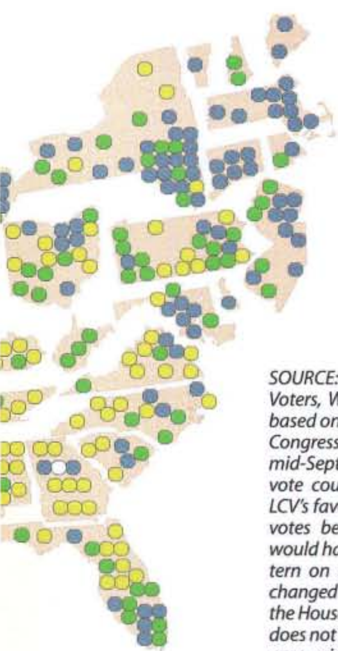
Conservatives feel that jobs should take priority over environmental legislation (in this respect they are supported by some unions). They feel that environmental regulations can be economically detrimental, a favorite example being the proposed Kyoto Treaty on global warming, which sets emission limits on human-made greenhouse gases. They believe the treaty could put the U.S. at a severe trade disadvantage and warn against treaties that give precedence to international law over U.S. law. For their part, environmentalists support regulation as a way of promoting the quality of life and preserving biodiversity. They see little danger to the economy from regulation and indeed see it as an economic stimulant. For example, some claim that the Kyoto Treaty would promote investment in new, nonpolluting energy technology.

Attitudes toward environmental legislation are clearly apparent in party voting patterns; the average Republican score in the House is 24, compared with 72 for Democrats. In the Senate, party differences are even greater: Republicans there average 12 and Democrats 86. But some Republicans scored 100 or slightly below, whereas some Democrats scored near 0. Republican average scores are highest in the Northeast, and Democratic average scores are high everywhere except among Southern House members. Overall, House and Senate averages for the 105th Congress to date are, respectively, 47 and 45, or roughly the same as they have been since 1970, when the LCV first began keeping records.

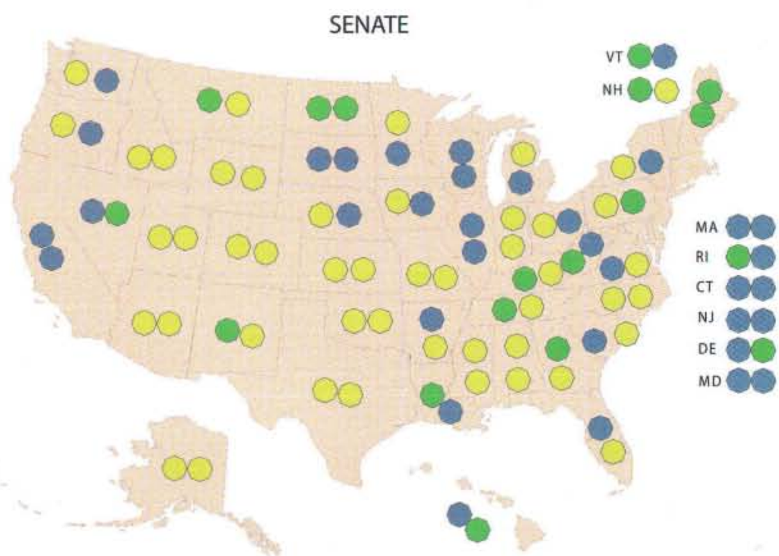
As of mid-September, few environmental bills of much substance had been passed by the 105th Congress. Among those that did are the Tropical Forest Conservation Act, which gives certain developing countries a financial incentive to spend more money on endangered habitats, and the National Parks and Wildlife Refuges Act, which clarifies the mission of these organizations. In the House, 140 members signed a letter supporting the Environmental Protection Agency's new and tougher clean air standards of 1997, thereby dampening any potential attempt by Congress to reverse the new standards. The Kyoto Treaty, which stands as one of the most far-reaching measures of the decade, had not been submitted to the Senate by press time, reportedly because President Bill Clinton did not feel that he could get the necessary two-thirds approval.

The scores of individual members of Congress and a descriptive listing of the measures on which the scores are based can be found at www.lcv.org on the League of Conservation Voters's World Wide Web site. Complete information on congressional voting records and debates can be located at thomas.loc.gov/-home/r105query.html on the Congressional Record Web site.

—*Rodger Doyle (rdoyle2@aol.com)*



SOURCE: League of Conservation Voters, Washington, D.C. Scores are based on voting records of the 105th Congress between January 1997 and mid-September 1998. A failure to vote counts as a vote against the LCV's favored position. Had the nonvotes been excluded, some scores would have been higher, but the pattern on the maps would not have changed substantially. The Speaker of the House (Newt Gingrich of Georgia) does not ordinarily vote. In the House map, circles indicate approximate position of the district in the state.



RODGER DOYLE

In Brief, continued from page 18

Quantum Error Correction

The promise of quantum computing—which relies on storing information in quantum states, such as the energy level or nuclear spin of an atom or molecule—has long been plagued by the problem of error correction. Quantum states are readily corrupted. But now researchers from Los Alamos National Laboratory and the Massachusetts Institute of Technology have devised a solution. Using radio-frequency pulses, they spread a single bit of quantum information onto three nuclear spins (rather than the conventional single bit onto a single spin) in individual molecules of trichloroethylene molecules in solution. In doing so, they have made it three times harder to introduce errors into the information.

Sprawling Suburbia

Strip malls, parking lots, housing developments and the like are eating up precious American farmland, according to a new report from the Sierra Club. In Atlanta the outward creep of urban development claimed an estimated 500 acres of farmland and forest a week last year. And Chicago's metropolitan area bloated by 40 percent between 1990 and 1996. The good news is that the U.S. Department of Agriculture recently pledged $17.2 million to protect productive farmland from commercialization.

CATHLYN MELLOAN *Tony Stone Images*

Zinc and Anorexia

Researchers led by Neil F. Shay of the University of Illinois have discovered that a diet poor in zinc may exacerbate anorexia nervosa, a condition in which patients starve themselves, occasionally to death. In articles published in *Nutritional Biology* and the *Journal of Nutrition,* the team reported that zinc deficiency in rats produced a rise in levels of neuropeptide Y—a compound that stimulates appetite in the brain. Shay speculates that although more neuropeptide Y is produced during zinc deficiency, its normal effect must somehow be dampened. The finding suggests that zinc supplements may help anorexics regain needed body weight.

—Kristin Leutwyler

ANTI GRAVITY

Lucky Laima

A television show from the 1970s featured an exceedingly diminutive fellow in evening dress who informed his boss at the start of every episode as to the impending arrival of their guests by hollering, "The plane! The plane!" The recent achievement of another aircraft conjured up the image of that obstreperous raconteur, as perhaps only he would have been small enough to pilot it. His presence, however, would have defeated the flight's purpose: the first transatlantic crossing by an unmanned airplane.

The robotic plane, one of a fleet called Aerosondes, is two meters (six feet) long, has a three-meter wingspan and weighs about as much as Herve Villechaize, tipping the scales at a wispy 13 kilograms (29 pounds) or so, depending on how much of its eight liters (two gallons) of fuel is left. The product of an outfit called the Insitu Group in Bingen, Wash., in conjunction with the University of Washington and an Australian group, the Aerosonde, dubbed Laima, departed Newfoundland on August 20. Although it left from an airport at Bell Island, that site was a formality rather than a necessity. The plane actually took off from the top of a speeding car, a launch strategy usually reserved for forgotten bags of groceries.

Computers tracked its progress for about 40 kilometers before Laima flew out of communication range. At that point, an onboard global positioning satellite navigation system guided Laima along a programmed route to Scotland's Outer Hebrides islands. Two other Aerosondes had failed to complete the flight earlier that week, so when engineers in Scotland reestablished contact with the craft about 25 hours after they last heard from it, they reenacted one of those Tranquillity-base-here-the-*Eagle*-has-landed, NASA-flight-control jubilation scenes.

Laima is the Latvian goddess of good fortune, and the name was a homage to the heritage of Juris Vagners, an aeronautics professor at the University of Washington and one of the project's key players. (Of course, people of science do not believe in appeals to mythological beings, but they know that such petitions sometimes work whether they believe in them or not.)

Transatlantic crossings have always represented watershed events. Failed attempts sometimes capture the imagination more than successes: witness the enduring *Titanic* frenzy and the relative inattention paid to Laima's safe landing. The human factor, obviously, has something to do with that. But maybe it's time to think of engineers in the heroic terms usually reserved for Lindberghs. Some lonely pilot's solo transatlantic excursion would have little value today, just as nothing that Edmund Hillary and Tenzing Norgay didn't see almost half a century ago awaits today's seemingly endless supply of Everest adventurers (other than all those discarded oxygen canisters).

Laima's quiet exploit could have actual significance. The oceanic crossing was designed to illustrate the potential of small, robotic aircraft, especially for weather reconnaissance on the other side of the continent.

Easterners are spoiled by high-quality forecasting, thanks to weather's propensity to move west to east over a vast landmass strewn with information-gathering stations. West Coasters, however, make do with much spottier forecasts because of the mostly instrument-free zone of ocean space. Although balloons, weather buoys and satellite images help, a small fleet of robot planes launched from Hawaii and Alaska, transmitting weather information on their way to the West Coast, could supply a whole lot of currently unavailable data points. They could also solve one of aviation's most vexing problems: nobody on board means no lost luggage.

—Steve Mirsky

MICHAEL CRAWFORD

PROFILE

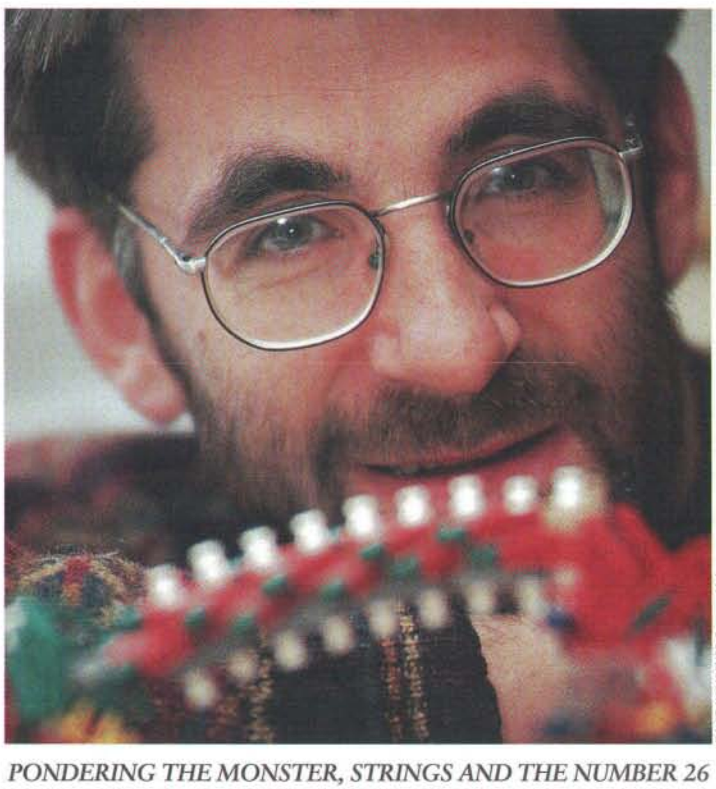
Monstrous Moonshine Is True

Richard Borcherds proved it—and discovered spooky connections between the smallest objects imagined by physics and one of the most complex objects known to mathematics

Talking to Richard Borcherds about his work can be unnerving. It is not just the difficulty of trying to keep up with the intellect of someone who, at the age of 38, has already won the highest award in mathematics, a Fields Medal, made of solid gold and bearing a Latin inscription that urges him "to transcend human limitations and grasp the universe."

There is also the palpable unease of his movements. I arrive at his office in a nondescript corner of the University of Cambridge precisely when he expected; I knock quietly on the door. Yet my entrance has completely flustered him. He begins pacing like a caged tiger and waving his arms at nothing in particular. He appears to have no idea what to do next. I offer myself a seat.

"I'm not very good at expressing feelings and things like that," Borcherds says straightaway. "I once read somewhere that the left side of the brain handles mathematics and the right side handles emotions and expression. And I've often had the feeling that there really is a disconnect of some sort between the two."


DAVID KAMPFNER *Gamma Liaison*

PONDERING THE MONSTER, STRINGS AND THE NUMBER 26 earned Richard Borcherds the highest honor in mathematics.

Mathematics research is not, as many believe, an exercise in pure reason—at least not for Borcherds. "The logical progression comes only right at the end, and it is in fact quite tiresome to check that all the details really work," he says. "Before that, you have to fit everything together by a lot of experimentation, guesswork and intuition."

That hints at what is most unnerving about talking to Borcherds: looking through his eyes, through his work, you can get a glimpse of a whole alternative universe, full of wondrous objects that are real but not physical. Borcherds spends his days exploring that deep space of mathematics, and indeed—if his frequent far-off stares and his choice today to dress entirely in wrinkled brown attire are any indication—he seems always to keep one foot over there.

"Some mathematics clearly is a human invention," he says, most notably anything that depends on the fact that we use a 10-digit numbering system. "But I think some mathematics does exist before its discovery. Take the Pythagorean theorem. That has been independently rediscovered several times by various civilizations. It's really there. Presumably if there were small furry creatures doing mathematics on Alpha Centauri prime, they would also have some version of the Pythagorean theorem."

And if they had explored a good deal further into the abstract universe of mathematics, the furry aliens might also have stumbled on three remarkable objects and discovered, as Borcherds did, that they are connected in some profound but still rather mysterious way. They would probably not, however, have called the problem the "monstrous moonshine conjecture," as Borcherds's mentor, Cambridge professor John H. Conway, chose to.

The problem arose in 1978, when John McKay of Concordia University was struck by a rather bizarre coincidence. "I was reading a 19th-century book on elliptic modular functions," McKay recalls, "and I noticed something strange in the expansion of one in particular"—the so-called *j* function. This elliptic modular function, explains John C. Baez of the University of California at Riverside, "shows up when you start studying the surfaces of doughnuts that are created by curling up the complex plane." On a sheet of graph paper, you can number the columns with whole numbers (1, 2, 3,...) and the rows with imaginary numbers ($1\sqrt{-1}$, $2\sqrt{-1}$, $3\sqrt{-1}$, . . .). Then you can roll up the sheet and join the ends of the tube to make doughnuts of various sizes and shapes. "Roughly speaking," Baez elaborates, "if you give me a shape of such a torus, then I can use the *j* function to convert that shape into a particular complex number." Although the *j* function sounds arcane, it is actually a useful tool in math and physics.

The odd coincidence appeared to McKay when he looked at the coefficients of the *j* function when it was written as an infinitely long sum. The third coefficient was 196,884. The number rang a bell.

To show me why, Borcherds lifts, with some effort, a book the size of a world atlas from his desk. He opens it to a table of numbers printed so small that they are barely legible. The first number in the table is 1. The next is 196,883. Together they add up to that figure in

Safety need not be for the lucky few.

At Opel, we consider safety to be a right, rather than a privilege. No matter how much one can afford to spend on a car.

It's the reason we designed the new Astra with our patented Pedal Release System. In the very unfortunate event of a full-frontal collision, the foot pedals on the new Astra are designed to detach, to reduce danger to the feet and legs. An industry first in this car segment.

At Opel, we don't believe in leaving anything to luck, especially when it comes to your safety.

Internet: http://www.opel.com

OPEL

the *j* function, which is mighty strange, because this table has nothing to do with elliptic functions. "These numbers," Borcherds says, flipping through about eight large pages of tiny print, "make up the character table of the Monster."

The Monster simple group is its full name, because it is the largest sporadic, finite, simple group known to exist. To understand what that means, Borcherds suggests, "suppose an ancient Greek tried to understand the symmetries of ordinary solid objects. He discovered a cube, which is quite easy to construct, and found that it has 24 symmetries"—that is, there are 24 ways to twist it about and end up with it looking the same. Those symmetries make up a finite group of 24 elements.

"Next, perhaps the Greek built a tetrahedron, which generates a group of 12 symmetries," Borcherds continues. "And then he might notice that no geometric object he knew of had a number of symmetries that is a multiple of five—but he could theorize that such an object exists. Later, somebody else might actually construct a [12-sided] dodecahedron, having 60 symmetries, thus proving the first guy right. In fact, it's said that when the Pythagoreans did discover a dodecahedron, they guarded it as such a great secret that they actually strangled one of their members who dared publicize its existence. They took their math seriously in those days."

Math rarely leads to murder anymore, but quite a few mathematicians have devoted the better part of their careers to solving the mysteries of the Monster group, which was indeed predicted to exist many years before it was successfully constructed. It, too, represents the symmetries of—well, of what exactly, mathematicians hadn't a clue. Something, certainly, that is a bit too complex to call a mere geometric object, because the Monster lives not in three dimensions but in 196,883. And in 21,296,876 dimensions, and in all the higher dimensions listed in the first column of Borcherds's table.

Whatever object gives rise to the Monster group must be exceedingly symmetrical, because "the group has several times more elements," McKay reckons, "than the number of elementary particles—quarks and electrons and such—in the sun": 808,017,424,794,-512,875,886,459,904,961,710,757,-005,754,368,000,000,000, to be precise.

So far removed are finite groups from modular functions that "when John McKay told people about his observation that the third coefficient of the *j* function matched the smallest dimensions of the Monster, they told him that he was completely crazy," Borcherds recounts. "There was no connection that anyone could imagine." But eventually others noticed that the coincidences ran too deep to ignore. "It turned out that *every* coefficient of the modular function is a simple sum of the numbers in this list of dimensions in which the Monster lives," Borcherds says. Conway and others theorized that the connections were not coincidences at all but reflections of some deeper unity. They dubbed the wild conjecture "moonshine," and a new specialty in mathematics arose to try to prove it.

Borcherds, meanwhile, was barely scraping through a Ph.D. at Cambridge, polishing his Go game when he was supposed to be at lectures, he says. Yet somehow he impressed Conway, who handpicked him to tackle moonshine.

In 1989, after eight years' work on the problem, Borcherds's body was sitting on a stalled bus in Kashmir while his mind no doubt roamed the alternative universe of the abstract. It was then that he found the third piece of the puzzle, the one that joined the other two. The connection, appropriately enough, was string theory, by way of the number 26.

Physicists have been dreaming up various kinds of infinitesimal strings for years in the hope of explaining everything in the universe with one theory. The basic idea is that all elementary particles are not fundamental at all but are really composed of loops of one-dimensional strings.

To keep track of how the laws of nature operate under various theories, physicists have long drawn stick-figure diagrams. Each limb represents the track of a particle, and the intersections, or vertices, are where the particles collide or interact. "In string theory they deal with little loops, not points, so the diagrams are made up not of lines but of tubes connected by bits of plumbing," Baez explains. "The math used in string theories describes what happens where these tubes meet," using a so-called vertex algebra.

One unpleasant fact about string theories, Baez notes, is that "when you try to do calculations in them, you need certain things to cancel each other out, but that only happens when you have 24 extra dimensions around," for a total of 26 dimensions (because time and the string itself take up two). "That is bad news for physicists," Borcherds says. "But it is exactly what you need to deal with the Monster. If the critical dimension of string theory were anything other than 26, I couldn't have proved the moonshine conjectures."

But it is, and he did, by inventing a vertex algebra—essentially, the rules of a string theory—all his own. "This vertex algebra," Baez explains, "describes a string wiggling around in a 26-dimensional space that has the unique feature that all 26 dimensions are curled up. It's like a tiny doughnut folded onto itself

***SYMMETRIES OF GEOMETRIC OBJECTS** and other mathematical constructs form the elements of so-called finite groups. A particular string theory, when applied to a folded doughnut in 26 dimensions (simplified to three here), has more than 10^{53} symmetries and produces the Monster group.*

in the coolest way, using a technique that only works in 26 dimensions."

Complex doughnuts, of course, are what the *j* function is all about. And, Borcherds showed, the Monster is simply the group of all the symmetries of this particular string theory—a theory, by the way, that almost certainly has nothing to do with the universe we live in. But it is now a well-explored land in the alternative universe that Borcherds spends most of his time in.

"There are a whole bunch of very spooky coincidences sitting together that get this to work," Baez reflects. "My feeling is that probably there is something even larger and deeper beneath it, something that hasn't been found. Borcherds has begun to uncover it. But there are still a lot of mysteries left." —*W. Wayt Gibbs in Cambridge, England*

CHEMICAL WEAPONS

PATENT BLUNDER

Terrorists' recipe for making the nerve agent VX in Sudan apparently came from a U.S. patent

Hours after American cruise missiles demolished a chemical plant in Sudan this past August, U.S. officials found themselves addressing Sudanese claims that the factory manufactured only pharmaceuticals and other beneficial compounds. The U.S., attempting to lend credence to its contention that the facility was producing chemical weapons, cited a soil sample obtained clandestinely a few yards from the plant this past June. The sample contained a chemical known by the acronym EMPTA, whose only practical, large-scale industrial use is in the manufacture of an extremely deadly nerve agent known as VX. The officials also insisted that Iraqi scientists had helped set up the VX operation at the Sudanese plant, a claim they said they confirmed by means of intercepted telephone conversations. Beyond those disclosures, however, the U.S. revealed little of the large, fragmented and incomplete mosaic of intelligence information that in all likelihood precipitated the site's selection for bombing.

This reticence may have been partly linked to an embarrassing fact: the heart of Iraq's recipe for VX may very well have come from the U.S. Patent and Trademark Office.

Iraq's affinity for U.S. patents and other open technical literature was established in 1991, shortly after the war in the Persian Gulf. Inspectors charged with uncovering Iraq's sprawling nuclear weapons program found that Iraqi scientists and engineers, in their push for an atomic bomb, made use of tens of thousands of pages of public documents on enriching uranium to weapons grade and on other nuclear topics, most of it declassified in the U.S. in the 1950s.

In all probability, Iraqi military scientists followed a similar strategy in their pursuit of chemical weapons. For example, of the several ways of making VX, Iraq chose to synthesize it from EMPTA, which stands for O-ethyl methylphosphonothioic acid. According to a U.S. intelligence source quoted in *Chemical & Engineering News*, Iraq and Sudan are the only countries to have taken the EMPTA route to VX. Three U.S. Army chemists developed the approach, which was the subject of a secret patent application in 1958. After being declassified in 1975, the patent became publicly available (it is now on the Internet). Joseph Epstein, one of the chemists who invented the EMPTA approach, believes it is "very possible that the patent has gotten into enemy hands."

First produced by British government chemists in the 1950s, VX is the most lethal of the four common military nerve agents. A single drop of the thick, amber-colored liquid on the skin can kill an adult. The Japanese terrorist organization Aum Shinrikyo produced some VX in a small laboratory near Mount Fuji and used the compound in Osaka in 1994 to kill a disillusioned former member of the sect. That 28-year-old man is the only known victim of VX.

Although it patented the EMPTA process, the U.S. never used it to produce VX in large quantities. According to Epstein, chemists developed other methods that were better suited to mass production.

Epstein speculates that Iraqi chemists favored the EMPTA method because of its simplicity—producing EMPTA, which is derived from phosphonic acid, is the only difficult part, he says. To then make VX, chemists need only mix the EMPTA and another reagent in room-temperature water and extract the nerve agent from the resulting solution.

Epstein also doubts that EMPTA would be used to make insecticides, as some reports have suggested, because the resulting products would be so highly toxic to mammals. He said it is possible that it could be used to make antimicrobial agents or fungicides, but he said such compounds would be relatively expensive to produce that way.

Epstein says he does not know why

AP PHOTO/RAOUF

***DEMOLISHED QUARTERS OF A SUDANESE PLANT** were said to be producing a deadly nerve agent.*

the U.S. government decided to declassify the patent. John F. Terapane, director of the licensing and review section of the U.S. Patent and Trademark Office, says secrecy orders are placed on, and removed from, patents on the recommendation of other government agencies—namely, the Department of Energy for nuclear weapons technologies and the Defense Technology Security Administration in the Department of Defense for essentially everything else. Peter Sullivan, deputy director of the Defense Technology Security Administration, declined to be interviewed for this article.

Frank Barnaby, a former nuclear weapons designer who writes and lec-

tures on national security, emphasizes the ubiquity of material on chemical weapons, from the *Merck Index* to the journal *Acta Chemica Scandinavica.* And according to *America the Vulnerable,* a 1987 book by Joseph D. Douglass and Neil C. Livingstone, a Swedish armed forces publication "describes in detail how to launch a gas attack, including formulae for calculating wind speed and lethal concentrations of the agent."

"If you're a good chemist, you've only got to know the chemical name of VX in order to guess a way of preparing it," Barnaby insists. "I suppose because of that, the body of chemists would see no reason for keeping [patent and other information] secret. There are no secrets about these things anymore." *—Glenn Zorpette, with additional reporting by Steven J. Frank, a patent attorney with Cesari and McKenna in Boston*

SOFTWARE

ROUGH SAILING FOR SMART SHIPS

Does commercial software such as Windows NT compromise naval ship performance?

Three years ago the U.S. Navy commenced a bold plan for slashing costs while preparing its fleet for the next century. The program, dubbed "Smart Ship," called for a reduction in crew levels through increasingly computerized ships. Additional savings would be achieved by using commercial off-the-shelf products, such as Pentium-chip computers, instead of expensive custom parts to build the new automated systems. But Smart Ship has recently encountered rough waters. A major computer crash on board the first of the automated ships has led to harsh criticisms of the navy initiative, and the dispute has touched off ugly accusations that important technical decisions are being controlled by politics—not by engineering.

BRIAN R. WOLFF IIPI

USS **YORKTOWN** *suffered a major computer crash in September last year.*

The controversy began when the USS *Yorktown,* a guided-missile cruiser that was the first to be outfitted with Smart Ship technology, suffered a widespread system failure off the coast of Virginia in September last year. After a crew member mistakenly entered a zero into the data field of an application, the computer system proceeded to divide another quantity by that zero. The operation caused a buffer overflow, in which data leak from a temporary storage space in memory, and the error eventually brought down the ship's propulsion system. The result: the *Yorktown* was dead in the water for more than two hours.

In a scathing article published in the June issue of the U.S. Naval Institute's *Proceedings,* Anthony DiGiorgio, an engineer for the Atlantic Fleet Technical Support Center, criticized the *Yorktown*'s deployment of Windows NT, a commercial operating system. "Using Windows NT ... on a warship is similar to hoping that luck will be in our favor," DiGiorgio wrote. Blaming NT for the *Yorktown*'s failure, some insiders groused that political pressures had forced the Microsoft operating system onto the ship.

Others insist that NT was not the culprit. According to Lieutenant Commander Roderick Fraser, who was the chief engineer on board the ship at the time of the incident, the fault was with certain applications that were developed by CAE Electronics in Leesburg, Va. As Harvey McKelvey, former director of navy programs for CAE, admits, "If you want to put a stick in anybody's eye, it should be in ours." But McKelvey adds that the crash would not have happened if the navy had been using a production version of the CAE software, which he asserts has safeguards to prevent the type of failure that occurred.

The mishap has provided ample ammunition to critics of Smart Ship, including contractors and navy staff whose livelihoods might be jeopardized by increasing reliance on commercial off-the-shelf (COTS) products, such as NT. "There's a faction in the navy that doesn't want Smart Ship to be successful," asserts Trey McKay of Intergraph, a supplier of Pentium-based PCs to the military. Indeed, Smart Ship upsets the cozy relationship between the Department of Defense and certain suppliers that have exacted premium prices for systems designed especially for the military.

For now, the navy's official stance remains unchanged. "We are absolutely committed to COTS ... and to the Windows NT operating system," insists Captain Charles Hamilton, deputy for Fleet in the Program Executive Office for Theater Surface Combatants. In fact, the navy has been trumpeting Smart Ship as a success, claiming that the *Yorktown* was able to reduce its crew by more than 10 percent, which could contribute to a potential annual savings of $2 million.

But other hurdles loom. The navy's plan to deploy Smart Ship technology on additional cruisers has been stalled by a protest filed by contractor Electronic Design, Inc. (EDI). The Government Accounting Office has recently upheld EDI's complaint, forcing the navy to revise its solicitation of bids for the next round of Smart Ship installations, which should have commenced earlier this year. *—Alden M. Hayashi*

GENE THERAPY

SHUTTING DOWN A GENE

Antisense drug wins approval

The idea seems simple and elegant. Turn off gene expression by blocking the action of the messenger RNA, which provides the essential information for assembling a protein. Antisense therapy, as it is called, could conceivably target a virus or a cancer cell with exquisite precision. But this biotechnology has followed the

trend line for much of the rest of the industry. Initial hyperbole was followed by disillusionment and even abandonment of the technology by some developers. Finally, a more balanced sense of realism emerged about future prospects.

"There was tremendous optimism among scientists and investors that these were going to be the drugs of the 1990s and the new millennium," remarks Arthur M. Krieg, a professor of internal medicine at the University of Iowa and an editor of an antisense journal. "It became clear very rapidly that things were not that easy."

In mid-August the U.S. Food and Drug Administration moved antisense therapeutics a modest step toward fulfilling some of its original promise. It approved fomivirsen (Vitravene), a drug made by Isis Pharmaceuticals in Carlsbad, Calif., that is injected into the eye to inhibit a viral infection in AIDS patients that can lead rapidly to blindness. The drug inhibits production of a protein that the virus needs to replicate. Because it specifically targets the viral RNA, it avoids some of the toxicity of other drugs used to treat cytomegalovirus retinitis. "Fomivirsen clearly demonstrates that you can make an antisense drug, translate it into a commercial reality, and it works," says Stanley T. Crooke, chairman and chief executive of Isis.

The first commercial antisense drug, which will be marketed by Ciba Vision, a unit of Novartis, is by no means the next blockbuster. It may garner revenues of little more than $50 million annually and perhaps much less—a pittance compared with the $1-billion-plus markets for Prozac and Viagra. The use of protease inhibitors—important components of AIDS cocktails—has reduced the patients who succumb to cytomegalovirus retinitis. In part because of the resulting small test population, the FDA approved the drug for patients who could not tolerate or were unresponsive to other treatments. Still, fomivirsen may pave the way for drugs that Isis and other companies have in their pipeline to fight maladies such as Crohn's disease, rheumatoid arthritis and cancer. "It's not going to make them profitable, but it certainly provides an income to allow them to learn more about the drug," says Steven P. Delco, senior biotechnology analyst for Miller Tabak Hirsch, a New York investment firm.

Building the antisense molecules—called oligonucleotides, the string of DNA that binds to a part of the messenger RNA—may have been an elegant idea, but it was by no means simple. Isis had to reengineer the oligonucleotides so that they are not immediately attacked by enzymes in the body that break down nucleic acids. But making certain changes to the nucleotide backbone can prevent it from attaching to the RNA. Antisense molecules, moreover, do not link to all sites along the messenger RNA, so researchers must undertake an extended process of trial and error to find just the right nucleotide sequences. The "oligos" can also interact with a cell in different ways, making it difficult to determine whether antisense binding is, in fact, producing the therapeutic effects.

More levelheaded expectations, it seems, have begun to emerge for antisense drugs. "It looks like it's a technology that works, but not in all organs and not for all indications," Krieg says. Muscle tissue does not easily take up antisense—and the molecules have difficulty crossing the blood-brain barrier. That still leaves the kidneys, liver and spleen, among other organs. And that may be enough for a radically new means of delivering drug therapy. —*Gary Stix*

Scientific American Resource Center

Welcome to Scientific American! For your convenience, we have summarized the answers to often-asked customer service questions and how to get in touch with us.

Scientific American Presents
Our new quarterly publication. Each 1998 issue is dedicated to a single theme: March, Magnificent Cosmos; June, Women's Health; September, Oceans; December, Intelligence. Annual subscriptions: In U.S.: $19.80. Elsewhere: $23.80 surface; $43.80 air service.

www.sciam.com
Our award-winning Internet resource, updated weekly and thoroughly linked to related sites. Here you will find timely and interesting current news, recent illustrated scientific articles, current issue highlights, ask the editors, our on-line store, subscriptions, e-mail, and other features not found in Scientific American, and much more.

Scientific American subscription rates
In U.S., Scientific American is: $34.97. Elsewhere: $49. Your 12-issue subscription includes all regular issues with the in-depth special reports. Our 1998 report subjects: April, Wireless Technologies; July, AIDS update; and October, Computer Security.

SCIENTIFIC AMERICAN

Subscription inquiries
Call us: In the U.S. and Canada: (800) 333-1199.
In other countries: (515) 247-7631.
e-mail: subscriptions@sciam.com
or write: Subscription Manager; Scientific American; PO Box 3187; Harlan, IA 51537.
The date of the last issue of your subscription appears on the mailing label. For change of address, please notify us at least four weeks in advance, and include both old and new addresses.

Back issues
In U.S. $9.95 each. Elsewhere: $12.95 each. Many issues are available. Fax your order with your Visa, MasterCard or AMEX information to: (212) 355-0408.

Reprints
$4 each, minimum order 10 copies, prepaid. Limited availability. Write to: Reprint Department; Scientific American; 415 Madison Avenue; New York, NY 10017-1111; fax: (212) 355-0408 or tel. (212) 451-8877.

Photocopying rights
Granted by Scientific American, Inc., to libraries and others registered with the Copyright Clearance Center (CCC) to photocopy articles in this issue for the fee of $3.50 per copy of each article plus $0.50 per page. Such clearance does not extend to photocopying articles for promotion or other commercial purpose.

MICROMECHANICS

LITTLE BANGS

Making thrusters for micromachines

Getting enough force out of silicon micromachines for them to do a useful amount of work has always proved a nettlesome challenge. A few researchers have begun to obtain more bang for the micron by making silicon chips with tiny cavities, filling them with explosives or rocket propellant and setting them afire.

Micropyrotechnics, it is conjectured, may one day power or reorient satellites and pulse drugs through the skin. Coupling ignitable materials with microelectromechanics (MEMS)—the technology that fashions submillimeter, electrically driven machines through standard chip fabrication methods—has begun to advance beyond the concept stage in a few laboratories.

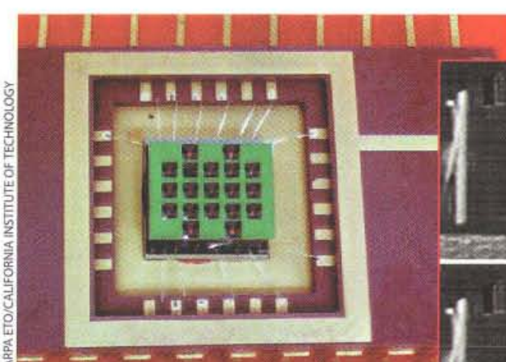
The Defense Advanced Research Projects Agency (DARPA) last year awarded a $3.5-million contract to TRW, Aerospace Corporation and the California Institute of Technology to come up with a prototype for a propulsion system that could be used to position or propel microsatellites for space, defense and communications applications. Micropyrotechnics takes advantage of the ability of silicon fabrication methods to produce lots of little devices at once.

The TRW-led team has so far built a chip that contains 15 thrusters—a five-by-three array of elements. A thruster is essentially a silicon box that measures about 700 to 1,000 microns on a side and is filled with a propellant such as lead styphnate.

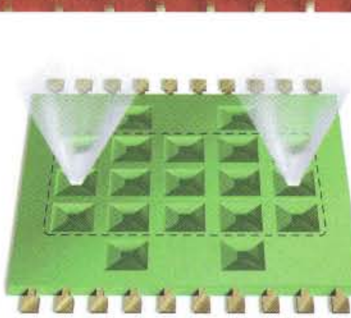
Each box has a microscopic electrical resistor that heats up when it receives a signal from control circuitry. This action lights the fuel, providing enough force to rupture one of the outer faces of the box, which is made thinner during manufacturing than the other side walls. A thruster element can be used only once, but arrays of thousands or millions of thrusters might keep a satellite going for a few years. The existing prototype, for instance, might be developed into a panel that would measure 100 square centimeters (almost 16 square inches) and contain roughly a million thrusters.

Adjusting propulsion in precise increments by lighting different numbers of thrusters has lent the technology the name "Digital Propulsion." "It's typically difficult to make engines have arbitrarily small units of thrust, but we can do that," says David H. Lewis, a TRW research engineer who invented the system with Erik K. Antonsson of Caltech. Microsatellites may measure as little as 10 centimeters along one edge, weigh one or two kilograms (up to 4.5 pounds) and be deployed from the space shuttle or a rocket.

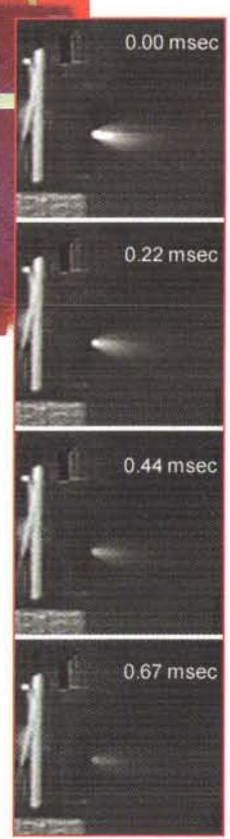
The National Aeronautics and Space Administration has considered them for space science. The Department of Defense is interested in them for use in ballistic-missile interceptors; a space-based projectile of less than a kilogram could accelerate to several kilometers per second, fast enough to obliterate a warhead traveling at an even higher speed. Communications companies could deploy clusters of thousands of satellites that together could function as a reconfigurable antenna whose position in space might change on command from a standard-size orbiting satellite.



DARPA ETO/CALIFORNIA INSTITUTE OF TECHNOLOGY



LAURIE GRACE



TINY POPS ARE EMITTED from microthrusters (surrounded by dashed line in diagram) *that were built by TRW and collaborators.*

Other work on micropyrotechnics continues at a French laboratory that has been involved with the technology for both space and medical applications. The Laboratory for the Analysis and Architecture of Systems (LAAS)/-CNRS has received a patent for a micropyrotechnic device that could replace a hypodermic needle or a transdermal patch. It consists of a microscopic hole in a silicon chip that might be filled with an explosive chemical, such as glycidyle azide polymer, which is used to set off air bags. Igniting the polymer would expand a silicon membrane. The movement of this membrane would send a volume of liquid at high velocity out of the device and through the skin. "If you consider an actual mechanical syringe, the speed of the injection is very slow by comparison," notes Carole Rossi, the LAAS researcher who developed the concept. "The duration of the pyrotechnic injection would be much shorter, and the pain would be diminished."

Micropyrotechnics could lead to still more ambitious schemes. Kristofer S. J. Pister of the Berkeley Sensor and Actuator Center at the University of California at Berkeley heads a research group that has begun DARPA-funded work on what he calls "smart dust." Investigators at Berkeley are fashioning small packages of temperature, pressure and other sensors that could be lifted for brief intervals by microthrusters to monitor weather or air quality or a battlefield. The sensors on these MEMS motes, each no more than a cubic millimeter in size, could then be interrogated by aircraft or unmanned aerial vehicles.

The Berkeley researchers, who have expanded on Rossi's work, want to send a smart dust particle a few hundred meters aloft using a single thruster. At its apex, the speck would deploy silicon wings coated with solar cells. The power generated could control the direction and rate of descent. Integrating sensors with electronics may let silicon chips see, hear and even smell. And adding micropropulsion will allow them to soar. *—Gary Stix*

CYBER VIEW

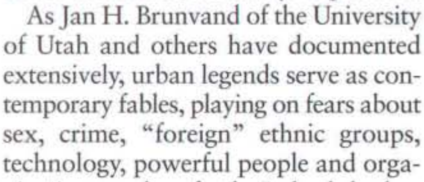
This Is Not a Hoax!

Help! Craig Furr, a six-year-old British boy with a brain tumor, wanted to go to Disney World before he died. After the trip, as his parents were checking out of the hotel, they noticed they'd been charged $2,500 for chocolate chip cookies from room service. While they were arguing over the bill, a gang of kidney thieves kidnapped Craig and spirited him away through Disney World's tunnels. The kidnappers put a wig on his head (which was hairless from chemotherapy) and dressed him as a girl, but luckily, a sharp-eyed security guard noticed Craig's old-fashioned side-lacing British shoes. The hotel, though, is suing Craig's parents for the $2,500, and unless they can get enough e-mail sent to santa@northpole.org in the next month to win the *Guinness Book of World Records*'s prize contest, they will lose their house!

You've never gotten this particular appeal in your electronic in box, but odds are that you've received—or sent—copies of at least some of the hundreds of urban legends circulating via the Internet. (In fact, this column was conceived when a *Scientific American* writer who shall remain nameless circulated the story of the $250 Neiman Marcus/Mrs. Fields/Waldorf-Astoria cookie recipe.) Temporary tattoos laced with LSD and strychnine, stolen kidneys, intimate encounters with gerbils, deadly computer viruses embedded in e-mail: the list goes on. Terry Chan of the Usenet discussion group alt.folklore.urban (AFU) maintains a file of more than 1,000 items of modern folklore, along with verdicts on their truth or falsity. You can read this FAQ (frequently asked questions) list or search for a verdict on a specific tale from www.urbanlegends.com and other sites.

But so few people do. As a result, the Internet appears to be at least as efficient at spreading myths as it is at disseminating truth. Urban legends used to spread by word of mouth, moving from city to city mostly with travelers, but now they can leap from one continent to the next in a few minutes. In addition, the rapid growth of the Internet provides an endless supply of "newbies," whose mental immune systems have not yet been toughened by exposure to hoaxes. Internet veterans recall the days when each new school year would bring a fresh crop of credulous correspondents, but now the influx is continuous, notes AFU regular and Internet manager Clive Feather. "1993 was the last September" that followed that pattern, he says—now, as far as the Net is concerned, it's always September.

As Jan H. Brunvand of the University of Utah and others have documented extensively, urban legends serve as contemporary fables, playing on fears about sex, crime, "foreign" ethnic groups, technology, powerful people and organizations and so forth. Indeed, he has traced some Internet tales to oral antecedents from the 1930s and before. There is, however, one big difference between word of mouth and word of e-mail: whereas oral traditions are almost



DAVID SUTER

always modified in the retelling, Internet legends can spread essentially unaltered. A few seconds' work with keyboard or mouse suffices to copy myths and forward them to a few thousand friends and neighbors. As the mythic culture goes global, it also becomes homogenized, Brunvand has lamented.

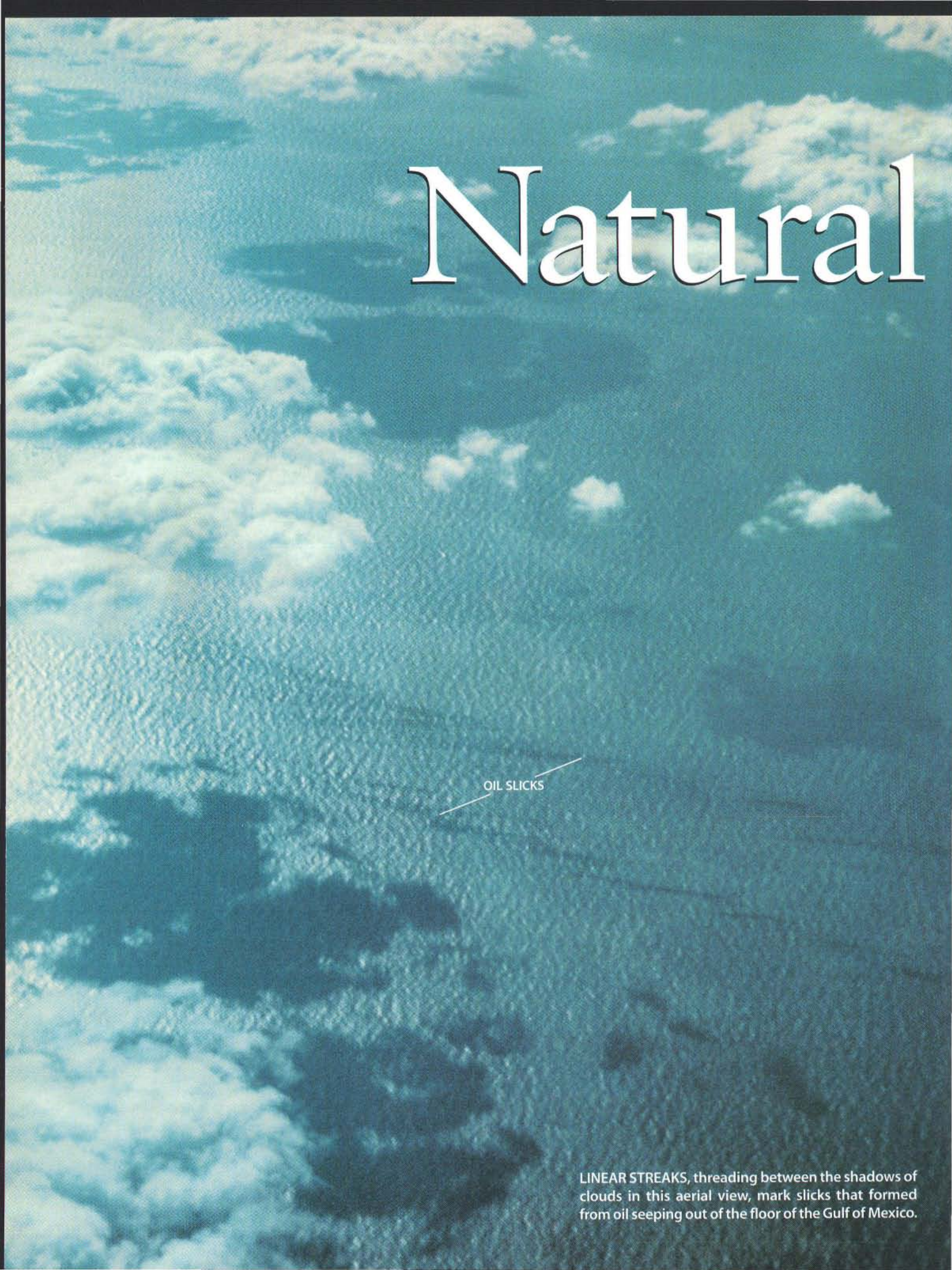
Like the "faxlore" that preceded them, Internet legends often contain some kind of call to action that helps them propagate (spread the cookie recipe, send a postcard, don't read a particular piece of e-mail, don't risk your kidneys by trysting with a beautiful stranger). Thanks to the essential untraceability of ASCII text, they may also have what seems like a solid provenance: it is simple to type "(AP)" or "(Reuters)" in front of a tale and make it look like wire-service copy. Of course, legitimate news outlets can also be taken in by urban legends, so even finding a story in a newspaper's archive or on its Web site is no guarantee of accuracy. At www.urbanlegends.com/medical/hospital_cleaning_lady.html or www.legends.org.za/arthur/cleanfaq.htm, you can read a thorough debunking of the legend about the cleaning woman who unplugged respirators to run her floor polisher, which was published by the South African newspaper the *Cape Times* and half a dozen others. Perhaps Internet legends will convince Net surfers to treat all the news they read with healthy skepticism.

Except in a few cases, it is usually impossible to determine the origin of Internet legends: they reappear every few months or years in slightly different versions, flood the virtual airwaves and then disappear. AFU regular Lee Rudolph says some newsgroup participants claim legends pop up on the Net after appearing in movies or television shows, but his own (equally anecdotal) observations don't bear the notion out. Each legend appears to have a characteristic period. "It would be nice," he writes, "to know what, if any, external forces either drive these periodicities or can overcome them to cause atypical sporadic outbreaks." Perhaps there is a hidden reservoir of the credulous, much like the isolated communities or populations of animal carriers that epidemiologists posit to explain the sporadic outbreak of dengue or flu.

Any single observer, of course, is ill placed to figure out how urban legends are propagating, because a legend could be waxing and waning Net-wide or just among a particular group of correspondents. Indeed, as Feather notes, Internet folklore aficionados might be worst placed of all because, for example, anyone who reads or posts to AFU already knows that urban legends exist. (And debunkers of e-mail soon find that their friends stop sending them the stories needed to build up a solid data set.)

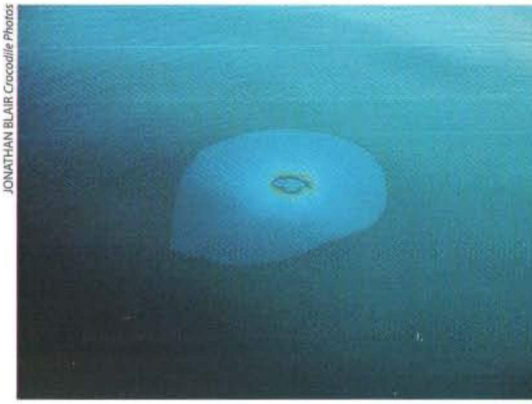
Until some Internet-wide folklore-monitoring system is in place, just keep sending those postcards to Craig Shergold (now a college student, he's been free of his brain tumor for more than seven years but still gets bags full of get-well mail thanks to the Net). And remember that as long as the Internet keeps growing, every day will feel like the start of the school year. *—Paul Wallich*

Natural



LINEAR STREAKS, threading between the shadows of clouds in this aerial view, mark slicks that formed from oil seeping out of the floor of the Gulf of Mexico.

Oil Spills



JONATHAN BLAIR Crocodile Photos

RAINBOW SHEEN is briefly visible when a drop of oil rises from depth, bursts through the surface and spreads. This image shows a plate-size patch.

In the Gulf of Mexico, a region famous for its many oil and gas fields, most of the petroleum flowing into the ocean leaks naturally from fissures in the seabed

by Ian R. MacDonald

Beneath the Gulf of Mexico, to the south of Texas and Louisiana, tiny bubbles of oil and natural gas trickle upward through faulted marine sediments. Close to the seafloor, these hydrocarbons ooze past a final layer teeming with exotic deep-sea life before they seep into the ocean above. Buoyant, they rise through the water in tight, curving plumes, eventually reaching the surface. There the gas merges with the atmosphere, and the oil drifts downwind, evaporating, mixing with water and finally dispersing.

The best time to witness such a natural "oil spill" is in summer, when the Gulf stays flat calm for days at a time. In the middle of the afternoon, with the full heat of the tropical sun blazing off the sea, one can stand on the deck of a ship and watch broad ribbons of oil stretch toward the horizon. Cruising upwind along one of these slicks, one will notice that the sea takes on an unusual smoothness. The clarity of the water seems to increase, and the glare of the sun off the surface intensifies. Flying fish break from the bow waves and plunge into the water again almost without making a splash. Presently, the scent of fresh petroleum becomes evident—an odor that is quite distinct from the diesel fumes wafting from the ship—and one sees waxy patches floating on the water or clinging to the hull.

Abruptly, droplets of oil begin bursting into little circles of rainbow sheen, which expand rapidly from the size of a saucer to a dinner plate to a pizza pan and then disappear, merging with the general glassy layer on the water and drifting away. Continue on an upwind course, and the sea regains its normal appearance. The water darkens, breezes can once again raise a tracery of tiny wavelets, and flying fish make their usual splashes. The ship has sailed beyond the oil slick. But off in the distance lies another and another. I have heard Coast Guard pilots say that before they knew better, they wasted hours flying up such slicks in search of a vessel spewing oil.

Indeed, the ongoing release of hydrocarbons from the seabed creates slicks that closely resemble the notorious results of surreptitious bilge pumping. Yet discharges of oil from the deep are a natural consequence of the geologic circumstances that make the Gulf of Mexico one of the great hydrocarbon basins of the world.

Time and Tide

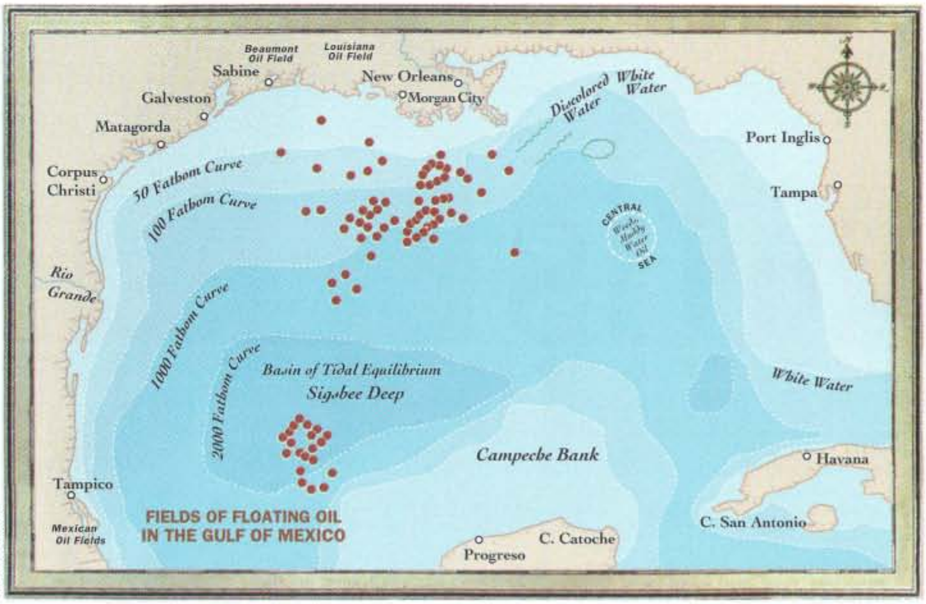
The oil that leaks upward from the bottom of the Gulf—like oil found everywhere—forms because geothermal energy constantly bakes the organic matter buried within sedimentary rock. Over time, the hydrocarbons created in this way rise from deeper layers until they become trapped in porous sandstones, fractured shales or the limestone remnants of ancient reefs.

Apart from having abundant source rocks and plentiful geologic "traps" for rising oil, the Gulf is also special because it contains an ancient salt bed, which was laid down during repeated episodes of evaporation in the Jurassic period, about 170 million years ago. This layer, known as the Louann Salt, underlies most oil fields in the region. The crystalline salt is malleable but relatively incompressible. Over geologic time, the weight of accumulated sediment—much of it transported offshore from land—has tended to force the salt upward and outward, forming sheets, spires or ridges. Some of these structures retain contact with the parent bed; others move as separate bodies through the surrounding sediment.

This so-called salt tectonism affects the migration of hydrocarbons in a number of ways. For example, salt is impermeable and can readily trap hydrocarbons below it. Also, the movement of salt can open large faults that extend from deeply buried reservoirs all the way to the surface, providing conduits through which petroleum can travel upward.

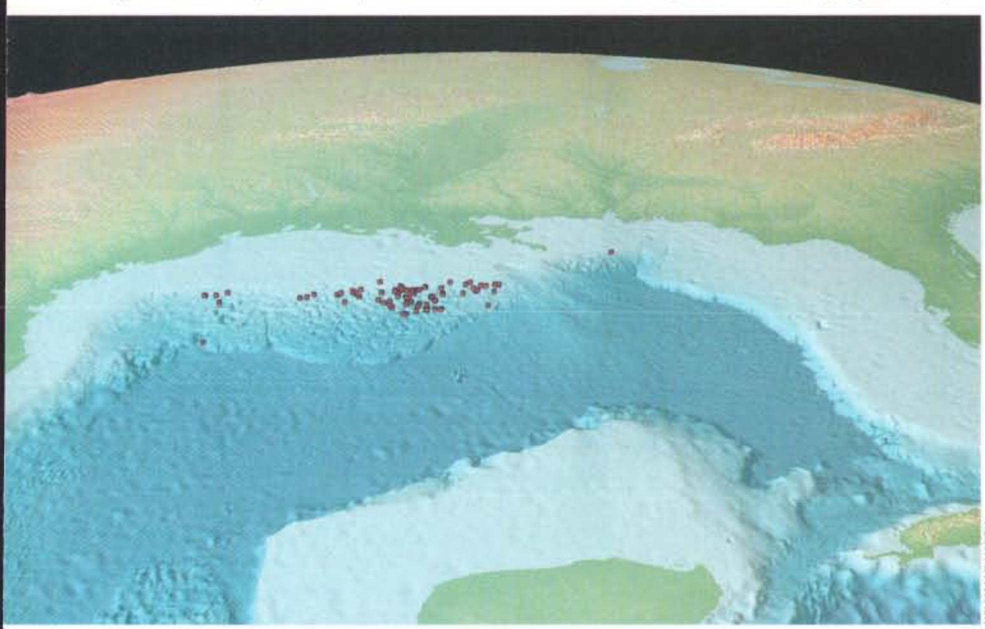
The presence of such structures makes the Gulf of Mexico a unique place, one

RALPH BAKER



JOHN DE SANTIS

SOLEY'S COMPILATION of oil slicks observed between 1902 and 1909 (*red dots above*) shows a preponderance of sightings southwest of the Mississippi Delta. The position of seafloor hydrocarbon seeps recently found in the Gulf of Mexico using modern methods (*dots below*) suggests that many of the seeps active at the turn of the century are still discharging oil today.



WILLIAM F. HAXBY

that looms very large in the history of the oil industry. Offshore oil and gas production was invented in the Gulf when the first platform was installed south of Louisiana in 1947. In subsequent years, operations moved farther and farther offshore as engineering advances made it possible to find and extract oil from below ever greater depths of water.

The social and economic consequences of this expansion have been pervasive. It is now almost impossible to imagine the Gulf coast without its population of oil workers, drilling rigs, production platforms, pipelines, tankers and refineries. To a degree not equaled elsewhere in the world, the Gulf of Mexico is a place where people live on and work under the sea—all to help satisfy society's insatiable hunger for petroleum.

Natural oil seeps offer a fascinating perspective on this enterprise. Although their existence comes as a surprise to many people, these seeps are well documented in the historical record. Pre-Columbian artifacts from the region show that tribal peoples commonly used beach tar as a caulking material, and Spanish records of floating oil date from the 16th century. In 1910 Lieutenant John C. Soley of the U.S. Navy published the first systematic study of offshore slicks in, curiously enough, *Scientific American* [see "The Oil Fields of the Gulf of Mexico," by John C. Soley; SCIENTIFIC AMERICAN SUPPLEMENT, No. 1788, April 9, 1910].

Soley reviewed sightings of oil at sea that had been noted in the handwritten logs of ships. A report from the steamship *Comedian*, for example, described oil coming up in three jets at one location. These accounts were written many decades before the first offshore platforms were built, at a time that even predates the widespread use of petroleum as a shipping fuel. So the source of all this oil was something of an enigma.

Soley's theories about the origin of the oil seem naive, in their details at least, in light of modern petroleum geology. For example, he thought that much of the petroleum found far from shore was inorganic. Still, the details recorded by these unnamed mariners—the floating mats of oil, fouled seabirds, lumps of tar and acrid stench of petroleum—evoke the all too familiar images of modern oil spills. In fact, the descriptions of turn-of-the-century oil slicks are so dramatic that one wonders whether the amount of seepage going on then was greater than it is today.

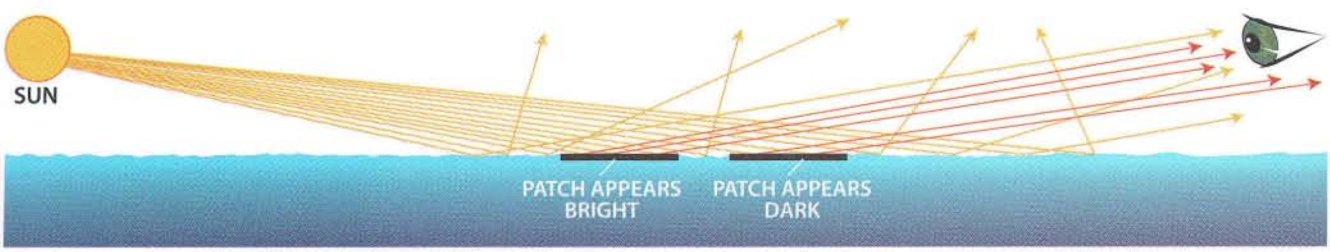
Seeking Seeps

Soley had to rely on anecdotal descriptions and imperfect navigation to map oil seeps. He probably could not have anticipated that the physical and chemical properties of floating oil make it visible from such great distances that scientists can now monitor it over vast areas.

When thicker than four microns—the magnitude that might result from a shipping accident—fresh oil forms an obvious covering, reddish brown to tan in color. With exposure, the volatile fraction of the oil rapidly evaporates, leaving a waxy residue that makes a foamy emulsion with seawater and tends to coagulate into gooey tar balls and floating mats. In thicknesses between one and four microns, an oily layer refracts incident light to form the rainbow sheen familiar from curbside puddles. Natural oil slicks—which range from less than 0.01 to one micron—may be only a few tens of molecules thick. Still, the chemical bonds between hydrocarbons are sufficient to create a surprisingly durable film. This surface-active,

SUN-GLINT PATTERN seen by astronauts in the space shuttle *Atlantis* in orbit over the Gulf of Mexico displays many separate oil slicks (*photograph at right*). Near the center of the glow, where the solar reflection is most intense, the slicks look especially bright. But at the periphery of the sun-glint pattern, the slicks appear dark. This difference arises because the mirrorlike surface of the slick reflects rays (*red lines in diagram below*) toward the viewer's eye in the first instance and away from it in the second, whereas the rougher surface of the unoiled water, which tends to scatter the incident light in both cases, appears more uniform.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION



JOHN DE SANTIS

or "surfactant," layer suppresses the fine-scale ripples that the wind would otherwise raise. The lack of ripples in turn allows the water to reflect light almost as effectively as a mirror, which gives a patch of floating oil its characteristic slick appearance.

The contrast between the reflection from a thin veneer of oil and the normal scattering of light from seawater makes these slicks quite distinctive under certain conditions. With the sun at a favorable angle, the slicks visible from a ship are also readily seen from airplanes and even from orbiting spacecraft.

Indeed, astronauts riding in the space shuttle can see slicks readily when they view the glint of the sun on the sea, a tactic they use to study ocean currents. In the center of a sun-glint pattern, the glare from oily patches is considerably brighter than the more diffuse reflection from unaffected waters. The situation reverses at the edges of the scene, where the geometry of illumination tends to direct light rays away from the viewer so that slicks look darker than their surroundings.

Slicks also appear dark under radar "illumination" because the source of the radio beam is in the same position as the detector, generally oblique to the water. The advent of satellite-mounted radars such as the European Radar Satellite, the Canadian RADARSAT and the space shuttle radar has meant that almost any location in the ocean can be monitored for traces of oil. Scientists sometimes use radar reflections from natural oil slicks to study how seawater circulates. The results they obtain can show details that would be difficult to discern with conventional oceanographic instruments. And geologists seeking telltale signs of hidden oil reservoirs can also take advantage of the views attainable from space.

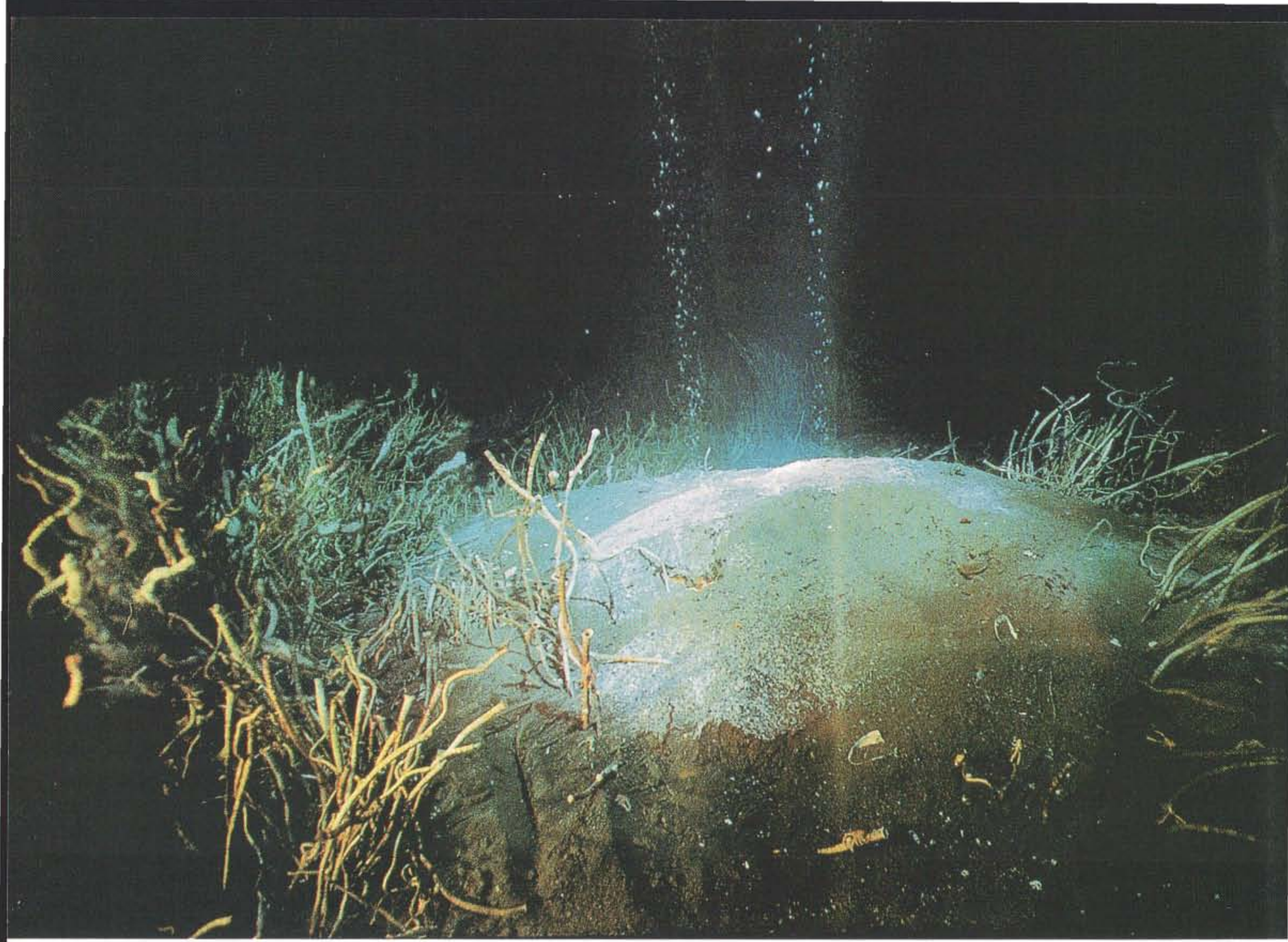
Sorting Wheat from Chaff

Using remote sensing to study natural oil seeps requires some way to distinguish the thin layers they generate from the vast majority of surfactant slicks, which have nothing to do with petroleum. Fish spawn, plant waxes and plankton swarms, among other phenomena, produce surfactants that, when concentrated by converging currents, can generate detectable slicks. But the markings of natural seepage are easily recognizable.

Typically, oil and gas leak from a group of vents spaced along a few hundred meters of a fault segment on the seafloor. Thus, the source is essentially fixed. Mathematical models indicate that a stream of oil drops rising from a single orifice through a kilometer of water will surface within a relatively small area. This theoretical result describes just what happens in in the Gulf of Mexico, where the rainbow traces of oil surfacing from a vent at the head of a slick occupy an area no more than 100 meters or so across. Over hours or days, currents at different depths will move this oily footprint around somewhat. But generally, the oil reaches the surface less than a kilometer or two from the underlying vent. So in practice, the repeated detection of a slick emanating from a stable location within an area of that size points to a source located on the seafloor.

This approach requires that one can tell the head of a slick from its tail. And here intuition is misleading. Familiar analogues, such as a plume of smoke rising from a cigarette, might suggest that the skinny end marks the origin. But the situation is fundamentally different. Whereas smoke particles are unlikely to reassemble themselves into a concentrated mass, spilled oil coagulates, forming the many individual bands that can be seen streaming away from a cluster of sources on the seabed.

At the start of each slick, oil drops expand after surfacing. Simultaneously, the oil drifts under the influence of the wind and the current. In principle, an oil drop floating on water should spread outward until it forms a layer that is only one molecule thick. In actu-



IAN MacDONALD

ality, the edges of a band of oil taper to the point where the slick cannot sustain itself. Downstream from the source, the layer becomes quite thin (on the order of 0.1 micron), but it is considerably thicker than the dimensions of a single molecule. Individual bands remain distinct for some distance, then merge and finally disappear. The overall outline of the slick is thus broader at its source than at its termination.

The length of a slick depends on the sea state. In heavy seas, an oily layer breaks up fairly rapidly and cannot drift far. Yet on a calm day, a layer of oil can be visible for 25 kilometers. Bends in the path traced by a slick reflect wind patterns: broad curves indicate gradual changes in wind direction, whereas chevrons document abrupt shifts. The time elapsed between a sudden reversal in wind direction and the moment a satellite acquires an image showing the resulting deviation indicates the life span of the slick. Using such comparisons, my co-workers at Texas A&M University and I typically find that the oil at the end of a visible slick has been floating on the surface for approximately 12 to 24 hours.

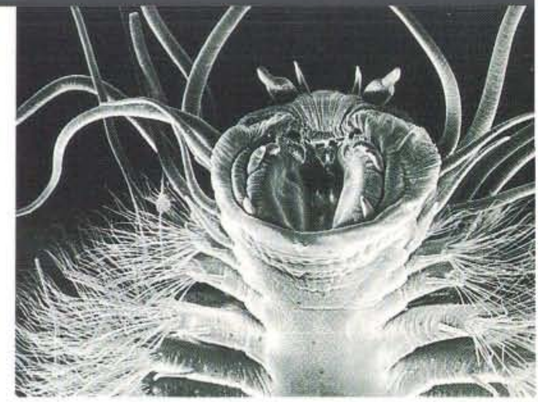
Living off Oil

Natural seeps may be a boon for oceanographers charting subtle eddies or for oil companies looking for new deposits to tap, but are they a bane for marine life? When the existence of the Gulf of Mexico oil seeps began to be widely recognized during the 1980s, my colleague Mahlon C. Kennicutt and his fellow researchers at Texas A&M speculated that the fauna living around seeps would provide a natural analogue for marine life exposed to oil pollution. To collect some of these presumably diseased specimens, they dragged a fishing net over active seeps. One of their first hauls, when it was brought on board their vessel, contained more than 800 kilograms of an unusual species of clam, *Calyptogena ponderosa*. Strangely, this large and obviously thriving creature was recovered from depths where deep-sea life normally proves rather scarce.

Adding to the mystery were dozens of brown, fibrous stalks also found in the net. These objects were so unfamiliar that the researchers almost threw them overboard, under the impression that the slender masses were merely some sort of reedy plant that had washed down the Mississippi River and settled in the deep Gulf. But one of the crew thought that this material might be good for weaving baskets. In sorting the stalks, he broke several of them open and spilled their red blood onto the deck, alerting the team that they had discovered something unexpected.

The specimens were eventually sent off to experts around the world. In the following months and years, a remarkable story began to emerge. Hydrocarbons leaking from the seafloor provide a source of chemical energy that nourishes creatures similar to the fauna first found at the hydrothermal vents of the Pacific Ocean in 1977. Vestimentiferan tube worms (the "stalks"), giant clams and a certain kind of deep-sea mussel

GAS-HYDRATE MOUND and a variety of deep-sea organisms surround a petroleum seep, seen here leaking gas from the floor of the Gulf of Mexico (*left*). At this site, the gas-hydrate material—a solid combination of natural gas and water that forms under great pressure—is home to a tiny animal dubbed the ice worm (*electron micrograph at right*). This remarkable polychaete worm was first uncovered by Charles R. Fisher of Pennsylvania State University during an expedition with the author last year.



CHARLES R. FISHER AND ROSEMARY WALSH *Pennsylvania State University*

make a living in both habitats through symbiosis with bacteria. Living within the cells of the animals, these bacteria synthesize new organic material in the absence of sunlight, using energy gained from oxidation of reduced compounds such as hydrogen sulfide.

The netting of vent creatures from the bottom of the Gulf of Mexico prompted investigators to take a closer look at this strange environment. In 1986 I led the expedition that made the first dives to an oil seep in a research submarine. My colleagues and I fully expected that we would have to search hard to find a few tube worms eking out a marginal existence. Instead we dropped right into the middle of a lush seafloor habitat, where we encountered large beds of mussels clustered around bubbling gas vents and extensive mats of brightly colored bacteria. Feeding on these exotic species was a diverse assemblage of fishes, crustaceans and other invertebrates that are commonly found in smaller numbers at shallower depths.

We now know that such thriving communities exist in many places in the Gulf. Interestingly, some of the biological activity at natural oil seeps tends to plug up the pores and fissures there. The metabolic by-products of microbes, in particular, cause the precipitation of calcium carbonate, which sometimes produces massive pavements that can trap oil below them. The creation of so-called gas hydrate on the ocean bottom can also block active gas vents.

Gas hydrate is an icelike substance that forms under conditions of high pressure and low but above-freezing temperature, when molecules of methane or other gases become trapped in a lattice of water molecules. Gas hydrates received early attention when these icy solids obstructed gas pipelines, forcing offshore operators to spend millions of dollars heating and insulating their undersea plumbing. More recently, researchers have focused on the gas hydrates that crystallize under the seafloor.

Nature's Pollution?

How much oil seeps naturally into the Gulf of Mexico? Assuming, conservatively, that an individual slick is 100 meters wide and maintains an average thickness of 0.1 micron over 10 kilometers, it must contain about 100 liters of oil. The life span of such a slick is typically 24 hours or less, indicating that while active, its source must have released at least 100 liters of oil a day. Estimating, conservatively again, that there are at least 100 such seeps in the Gulf at any time, then almost 40 million liters flow into the Gulf every decade.

It gives one pause to recall that the grounding of the *Exxon Valdez*, the benchmark of oil spills, dumped roughly the same amount of oil into Alaska's Prince William Sound. But it is staggering to consider that the releases in the Gulf of Mexico have been going on for a million years or more. Clearly, the ecosystem there has been able to cope with chronic oil "pollution" since long before the term was invented.

Comparing the natural release of oil through faults and fissures to its accidental release in the course of drilling or transport can be quite instructive—both for where the analogies hold and for where they break down. In the Gulf of Mexico, and probably in other parts of the world as well, natural leakage has extracted at least as much oil and gas from buried reservoirs as the petroleum industry has. But even though the total quantities may be roughly equivalent, the rates are not. Compared to nature, humankind is in a terrible hurry to get oil out of the ground.

That difference explains why a natural seep is not equivalent to a tanker accident, although the dosage might be identical. Just as a person who showers every day for a year may suffer the same exposure to water as one who drowns in a swimming pool, it is clear where the harm lies. The fact of natural oil seepage in no way forgives oil pollution. Whereas the open sea may tolerate hundreds of tons of oil a month—if it is spread thinly over tens of thousands of square kilometers—the same amount dumped on a seabird nesting area can obliterate the local population. Likewise, the chronic release of oil into an estuary may overwhelm that ecosystem.

Scientists and environmentalists alike must recognize that some oil spills can be quite damaging but that others are a benign part of the natural marine environment. The trick is to distinguish one from the other and to react appropriately. We do not want to be like the uninformed pilots, wasting time and energy just to chase sheen. SA

The Author

IAN R. MACDONALD began work in marine science as a fisheries development volunteer in Haiti. After working at the International Ocean Institute in Malta and at the United Nations Food and Agriculture Organization in both Malta and Rome, he returned to the U.S. and enrolled at Texas A&M University. There MacDonald obtained a master's degree in 1984 and a doctorate in 1990 after studying the ecology of the biological communities that flourish at natural oil seeps in the Gulf of Mexico. His current research at Texas A&M has also taken him recently to the Canadian Pacific and to the Caspian Sea.

Further Reading

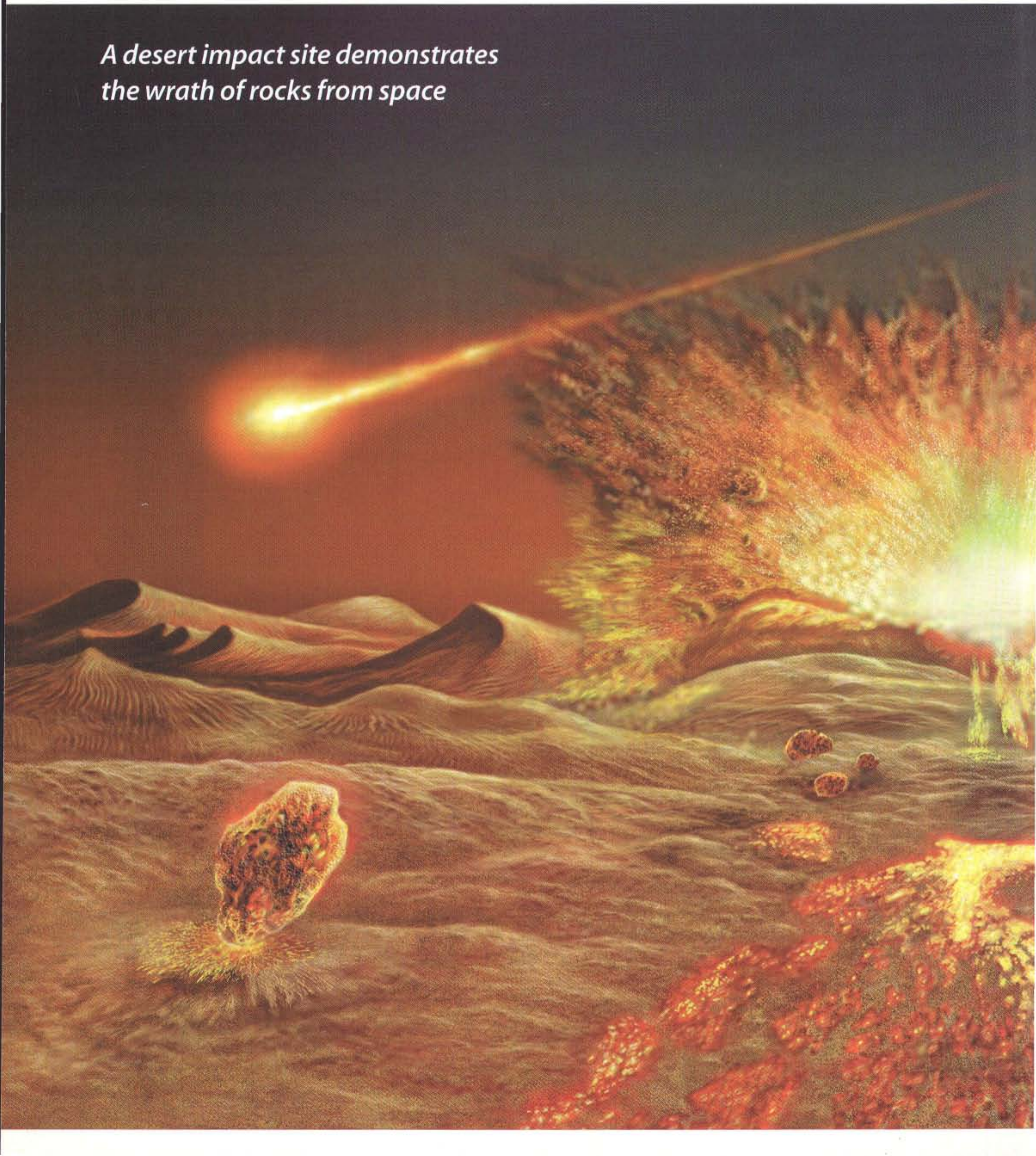
REASSESSMENT OF THE RATES AT WHICH OIL FROM NATURAL SOURCES ENTERS THE MARINE ENVIRONMENT. K. A. Kvenvolden in *Marine Environmental Research*, Vol. 10, pages 223–243; 1983.

NATURAL OIL SLICKS IN THE GULF OF MEXICO VISIBLE FROM SPACE. I. R. MacDonald et al. in *Journal of Geophysical Research*, Vol. 98, No. C9, pages 16351–16364; September 15, 1993.

REMOTE SENSING INVENTORY OF ACTIVE OIL SEEPS AND CHEMOSYNTHETIC COMMUNITIES IN THE NORTHERN GULF OF MEXICO. I. R. MacDonald et al. in *Hydrocarbon Migration and Its Near-Surface Expression*. Edited by D. Schumacher and M. A. Abrams. American Association of Petroleum Geologists Memoir 66, 1996.

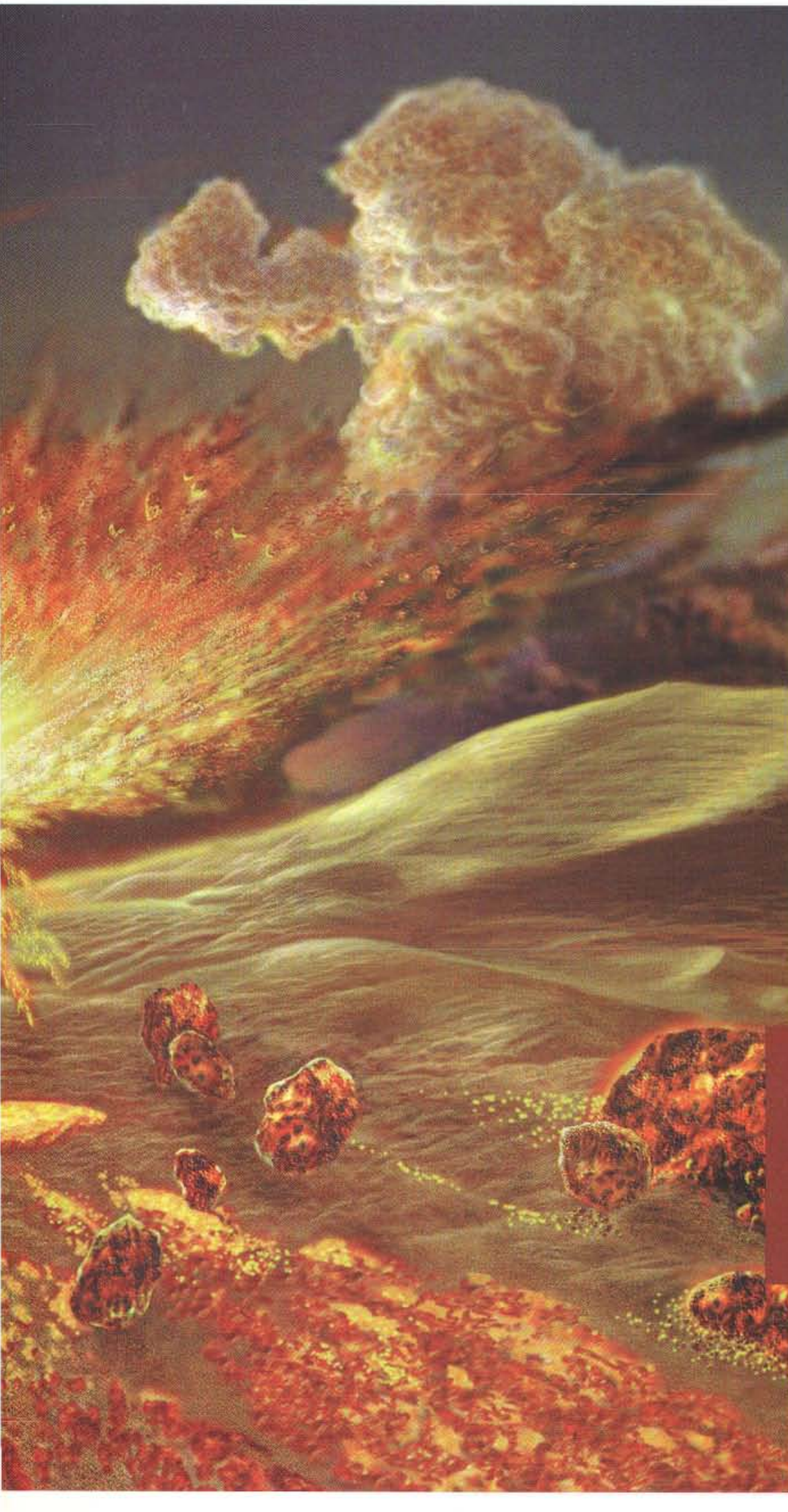
The Day the Sands Caught

A desert impact site demonstrates the wrath of rocks from space



Fire

by Jeffrey C. Wynn and Eugene M. Shoemaker



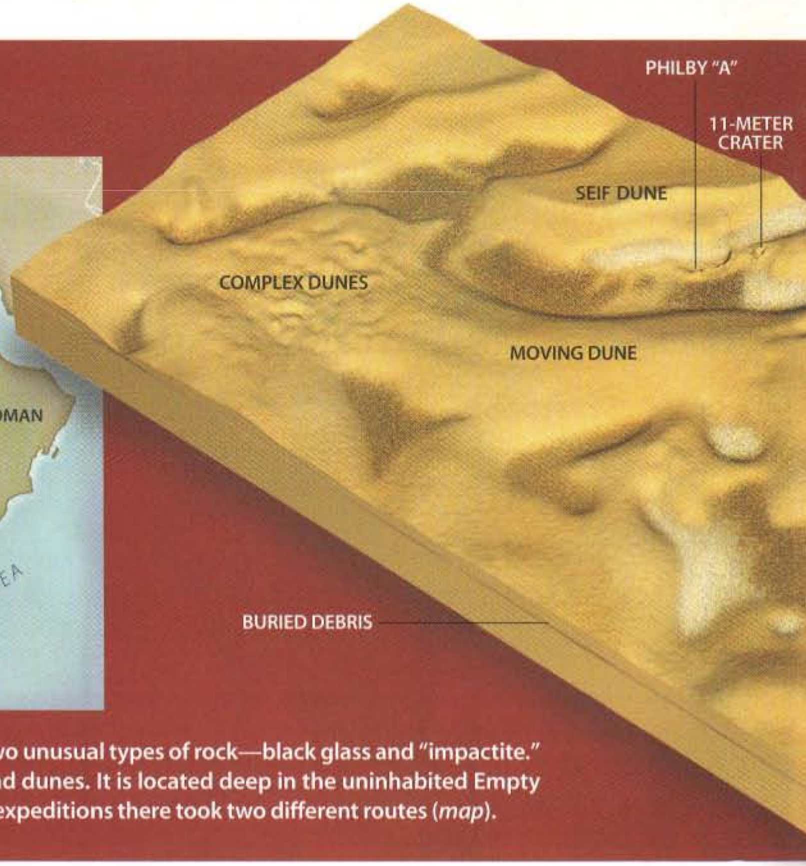
DON DIXON

Imagine, for a moment, that you are standing in the deep desert, looking northwest in the evening twilight. The landscape is absolutely desolate: vast, shifting dunes of grayish sand stretch uninterrupted in all directions. Not a rock is to be seen, and the nearest other human being is 250 kilometers away. Although the sun has set, the air is still rather warm—50 degrees Celsius—and the remnant of the afternoon sandstorm is still stinging your back. The prevailing wind is blowing from the south, as it always does in the early spring.

Suddenly, your attention is caught by a bright light above the darkening horizon. First a spark, it quickly brightens and splits into at least four individual streaks. Within a few seconds it has become a searing flash. Your clothes burst into flames. The bright objects flit silently over your head, followed a moment later by a deafening crack. The ground heaves, and a blast wave flings you forward half the length of a football field. Behind you, sheets of incandescent fire erupt into the evening sky and white boulders come flying through the air. Some crash into the surrounding sand; others are engulfed by fire.

Glowing fluid has coated the white boulders with a splatter that

The Wabar Meteorite Impact Site



WABAR SITE consists of three craters and a sprinkling of two unusual types of rock—black glass and "impactite." Much of the site has been buried by the ever shifting sand dunes. It is located deep in the uninhabited Empty Quarter of Saudi Arabia, or the Rub' al-Khali; the authors' expeditions there took two different routes (*map*).

first looks like white paint but then turns progressively yellow, orange, red and finally black as it solidifies—all within the few seconds it takes the rocks to hit the ground. Some pieces of the white rock are fully coated by this black stuff; they metamorphose into a frothy, glassy material so light that it could float on water, if there were any water around. A fiery mushroom cloud drifts over you now, carried by the southerly breeze, blazing rainbow colors magnificently. As solid rocks become froth and reddish-black molten glass rains down, you too become part of the spectacle—and not in a happy way.

Deep in the legendary Empty Quarter of Saudi Arabia—the Rub' al-Khali—lies a strange area, half a square kilometer (over 100 acres) in size, covered with black glass, white rock and iron shards. It was first described to the world in 1932 by Harry St. John "Abdullah" Philby, a British explorer perhaps better known as the father of the infamous Soviet double-agent Kim Philby. The site he depicted had been known to several generations of roving al-Murra Bedouin as *al-Hadida*, "the iron things."

There is a story in the Qur'an, the holy book of Islam, and in classical Arabic writings about an idolatrous king named Aad who scoffed at a prophet of God. For his impiety, the city of Ubar and all its inhabitants were destroyed by a dark cloud brought on the wings of a great wind. When Philby's travels took him to the forbidding Empty Quarter, his guides told him that they had actually seen the destroyed city and offered to take him there. Philby gladly accepted the offer to visit what he transliterated in his reports as "Wabar," the name that has stuck ever since.

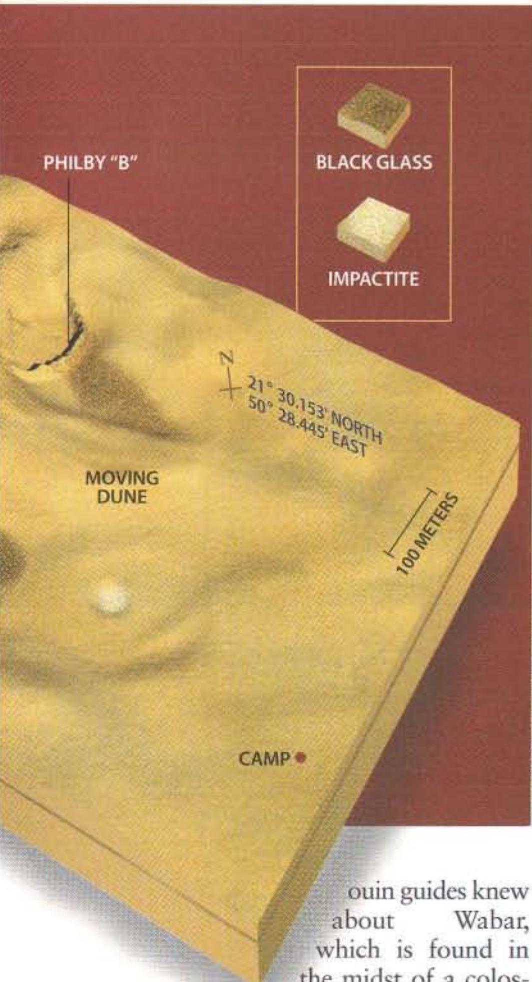
What he found was neither the lost city of Ubar nor the basis for the Qur'anic story. But it was certainly the setting of a cataclysm that came out of the skies: the arrival of a meteorite. Judging from the traces left behind, the crash would have been indistinguishable from a nuclear blast of about 12 kilotons, comparable to the Hiroshima bomb. It was not the worst impact to have scarred our planet over the ages. Yet Wabar holds a special place in meteor research. Nearly all known hits on the earth have taken place on solid rock or on rock covered by a thin veneer of soil or water. The Wabar impactor, in contrast, fell in the middle of the largest contiguous sand sea in the world. A dry, isolated place, it is perhaps the best-preserved and geologically simplest meteorite site in the world. Moreover, it is one of only 17 locations—out of a total of nearly 160 known impact structures—that still contain remains of the incoming body.

In three grueling expeditions to the middle of the desert, we have reconstructed the sequence of events at Wabar.

__A fiery mushroom cloud__ drifts over you now, carried by the southerly breeze. As solid rocks become froth and reddish-black molten glass rains down, you too become part of the spectacle—and not in a happy way.

The impact was an episode much repeated throughout the earth's geologic and biological history. And the solar system has not ceased to be a shooting gallery. Although the biggest meteors get most of the attention, at least from Hollywood, the more tangible threat to our cities comes from smaller objects, such as the one that produced Wabar. By studying Wabar and similarly unfortunate places, researchers can estimate how often such projectiles strike the earth. If we are being shot at, there is some consolation in knowing how often we are being shot at.

One has to wonder how Philby's Bed-



LAURIE GRACE (map); SLIM FILMS (illustration)

ouin guides knew about Wabar, which is found in the midst of a colossal dune field without any landmarks, in a landscape that changes almost daily. Even the famously tough desert trackers shy away from the dead core of the Empty Quarter. It took Philby almost a month to get there. Several camels died en route, and the rest were pushed to their limits. "They were a sorry sight indeed on arrival at Mecca on the ninetieth day, thin and humpless and mangy," Philby told a meeting of the Royal Geographical Society on his return to London in 1932.

Otherworldly

When he first laid eyes on the site, he had become only the second Westerner (after British explorer Bertram Thomas) to cross the Empty Quarter. He searched for human artifacts, for the remains of broken walls. His guides showed him black pearls littering the ground, which they said were the jewelry of the women of the destroyed city. But Philby was confused and disappointed. He saw only black slag, chunks of white sandstone and two partially buried circular depressions that suggested to him a volcano. One of his guides brought him a piece of iron the size of a rabbit. The work of the Old People? It slowly dawned on Philby that this rusty metal fragment was not from this world. Laboratory examination later showed that it was more than 90 percent iron, 3.5 to 5 percent nickel and four to six parts per million iridium—a so-called sidereal element only rarely found on the earth but common in meteorites.

The actual site of the city of Ubar, in southern Oman about 400 kilometers (250 miles) south of Philby's Wabar, was uncovered in 1992 with the help of satellite images [see "Space Age Archaeology," by Farouk El-Baz; Scientific American, August 1997]. Wabar, meanwhile, remained largely unexplored until our expeditions in May 1994, December 1994 and March 1995. The site had been visited at least twice since 1932 but never carefully surveyed.

It was not until our first trip that we realized why. One of us (Wynn) had tagged along on an excursion organized by Zahid Tractor Corporation, a Saudi dealer of the Hummer vehicle, the civilian version of the military Humvee. To promote sales of the vehicle, a group of Zahid managers, including Bill Chasteen and Wafa Zawawi, vowed to cross the Empty Quarter and invited the U.S. Geological Survey mission in Jeddah to send a scientist along. This was no weekend drive through the countryside; it was a major effort requiring special equipment and two months of planning. No one had ever crossed the Empty Quarter in the summer. If something went wrong, if a vehicle broke down, the caravan would be on its own: the long distance, high temperatures and irregular dunes preclude the use of rescue helicopters or fixed-wing aircraft.

An ordinary four-wheel-drive vehicle would take three to five days to navigate the 750 kilometers from Riyadh to Wabar [*see map on opposite page*]. It would bog down in the sand every 10 minutes or so, requiring the use of sand ladders and winches. A Hummer has the advantage of being able to change its tire pressure while running. Even so, the expedition drivers needed several days to learn how to get over dunes. With experience, the journey to Wabar takes a long 17 hours. The last several hours are spent crossing the dunes and must be driven in the dark, so that bumper-mounted halogen beams can scan for the unpredictable 15-meter sand cliffs.

Our first expedition stayed at the site for a scant four hours before moving on. By that time, only four of the six vehicles still had working air conditioners. Outside, the temperature was 61


JEFFREY C. WYNN

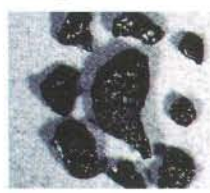
SAND-FILLED CRATER, 11 meters (36 feet) in diameter, was discovered by the authors on their expedition to Wabar in December 1994. Under the sand the crater is lined with a bizarre kind of rock—impactite—thought to have formed when immense pressures glued sand grains together. Around the crater rim are centimeter-size chips of iron and nickel. From the size of the crater geologists estimate that it formed when a dense metallic meteorite just one meter across smacked into the sand. This meteorite had split off from the larger bodies responsible for the other two craters at Wabar.

Identifying Impact Craters

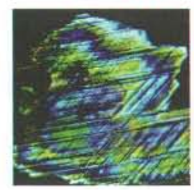
How would you recognize an impact crater if you fell into one? It isn't easy. Although the moon is covered with craters, it has no water, no weather, no continental drift—so the craters just stay where they formed, barely changed over the aeons. On the earth, however, all these factors have erased what would otherwise have been an equally pockmarked surface. To confuse matters further, more familiar processes—such as volcanism and erosion—also leave circular holes. Not until early this century did geologists first confirm that some craters are caused by meteorites. Even today there are only about 160 known impact structures.

Only about 2 percent of the asteroids floating around in the inner solar system are made of iron and nickel, whose fragments are fairly easy to recognize as foreign. But other types of meteorites blend in with the rest of the stones on the ground. The easiest place to pick them out is in Antarctica, because few other rocks find their way to the middle of an ice field. Elsewhere, recognizing a meteorite crater requires careful mapping and laboratory work. Geologists look for several distinctive features, which result from the enormous velocities and pressures involved in an impact. Even a volcanic eruption does not subject rocks to quite the same conditions. —*J.C.W.*

• **Shatter cones.** These impressions, found in the rocks around a crater, look like cookie-cutter cones or chevrons. Occasionally, you can see them in rock outcroppings if the cones have fractured lengthwise. No shatter cones appear at Wabar because the site formed in loose sand.



• **High-temperature rock types.** Laminated and welded blocks of sand have been seen at Wabar and at nuclear test sites. In addition, tektites—glassy rocks thought to form when molten rock is splattered into orbit and then solidifies on the way back down—appear around many large impact sites.



• **Microscopic rock deformation.** The crystal structure of some minerals is transformed by the shock waves during an impact. Quartz, for example, develops striations that are oriented in more than one direction. It can also recrystallize into new minerals, such as coesite and stishovite, detectable only in x-ray diffraction experiments.

CAROLYN SHOEMAKER

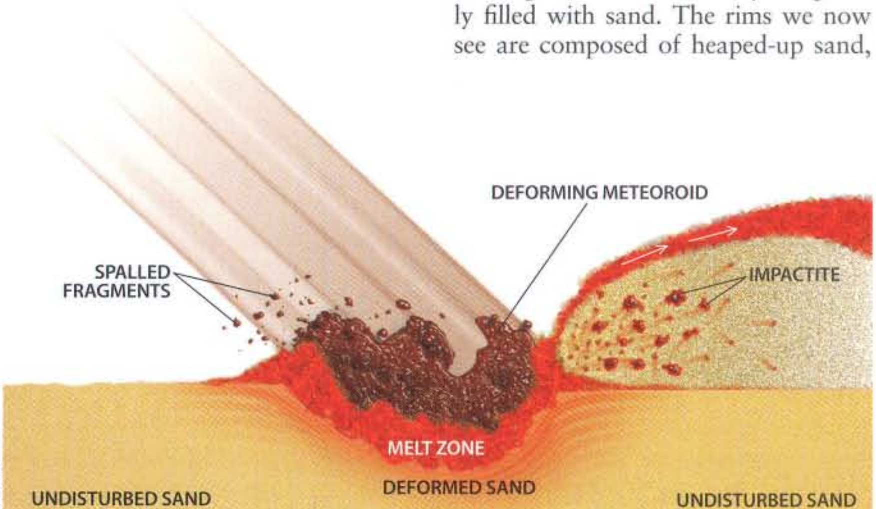
degrees C (142 degrees Fahrenheit)—in the shade under a tarp—and the humidity was 2 percent, a tenth of what the rest of the world calls dry. Wynn went out to do a geomagnetic survey, and by the time he returned he was staggering and speaking an incoherent mixture of Arabic and English. Only some time later, after water was poured on his head and cool air was blasted in his face, did his mind clear.

Zahid financed the second and third expeditions as well. On our weeklong third expedition, furious sandstorms destroyed our camp twice, and the temperature never dropped below 40 degrees C, even at night. We each kept a two-liter thermos by our beds; the burning in our throats awoke us every hour or so.

Shocking Rock

The Wabar site is about 500 by 1,000 meters in size. There are at least three craters, two (116 and 64 meters wide) recorded by Philby and the other (11 meters wide) by Wynn on our second expedition. All are nearly completely filled with sand. The rims we now see are composed of heaped-up sand, anchored in place both by "impactite" rock—a bleached, coarse sandstone—and by large quantities of black-glass slag and pellets. These sandy crater rims are easily damaged by tire tracks. There are also occasional iron-nickel fragments.

Geologists can deduce that a crater was produced by meteorite impact—rather than by other processes such as erosion or volcanism—by looking for signs that shock waves have passed through rocks [*see box above*]. The impactite rocks at the Wabar site pass the test. They are coarsely laminated, like other sandstones, but these laminations consist of welded sand interspersed with ribbonlike voids. Sometimes the



SLIM FILMS

CROSS SECTION of meteorite impact, as reconstructed by computer simulations, shows how the Wabar craters were created within a matter of seconds. The meteorite flattened as it hit the ground; a shock wave traveled backward through the body, causing part of it to spall off with little damage; the rest of the meteorite melted and amalgamated with sand directly underneath; surrounding sand was compressed into impactite. The whole mess was then thrust into the air. Deeper layers of sand were relatively unaffected.

layers all bend and twist in unison, unlike those in any other sandstone we have ever seen. The laminations are probably perpendicular to the path taken by a shock wave. Moreover, the impactite contains coesite, a form of shocked quartz found only at nuclear blast zones and meteorite sites. X-ray diffraction experiments show that coesite has an unusual crystal structure, symptomatic of having experienced enormous pressures.

The impactite is concentrated on the southeastern rims and is almost entirely absent on the north and west sides of the craters. This asymmetry suggests that the impact was oblique, with the incoming objects arriving from the northwest at an angle between 22 and 45 degrees from the horizontal.

The two other types of rock found at Wabar are also telltale signs of an impact. Iron-nickel fragments are practically unknown elsewhere in the desert, so they are probably remnants of the meteorite itself. The fragments come in two forms. When found beneath the sand, they are rusty, cracked balls up to 10 centimeters in diameter that crumble in the hand. Daniel M. Barringer, an American mining engineer who drilled for iron at Meteor Crater in Arizona early this century, called such fragments, which occur at several iron-meteor sites, "shale balls."

When the iron fragments are found at the surface, they are generally smooth, covered with a thin patina of black desert varnish. The largest piece of iron and nickel is the so-called Camel's Hump, recovered in a 1965 expedition and now displayed at King Saud University in Riyadh. This flattened, cone-shaped chunk, weighing 2,200 kilograms (2.43 tons), is probably a fragment that broke off the main meteoroid before impact. Because the surface area of an object is proportional to its radius squared, whereas mass is proportional to the radius cubed, a smaller object undergoes proportionately more air drag. Therefore, a splinter from the projectile slows down more than the main body; when it lands, it may bounce rather than blast out a crater.

The other distinctive type of rock at Wabar is the strange black glass. Glassy rock is often found at impact sites, where it is thought to form from molten blobs of material splattered out from the crater. Near the rims of the Wabar craters, the black glass looks superficially like Hawaiian pahoehoe, a ropy, wrinkled rock that develops as thickly flowing lava cools. Farther away, the glass pellets become smaller and more droplike. At a distance of 850 meters northwest of the nearest crater, the pellets are only a few millimeters across; if there are any pellets beyond this distance, sand dunes have covered them. When chemically analyzed, the glass is uniform in content: about 90 percent local sand and 10 percent iron and nickel. The iron and nickel appear as microscopic globules in a matrix of melted sand. Some of the glass is remarkably fine. We have found filigree glass-splatter so fragile that it does not survive transport from the site, no matter how well packaged.

The glass distribution indicates that the wind was blowing from the southeast at the time of impact. The wind direction in the northern Empty Quarter is seasonal. It blows from the north for 10 months of the year, sculpting the huge, horned barchan sand dunes. But during the early spring, the wind switches direction to come from the southeast. Spring is the desert sandstorm season that worried military planners during the Gulf War; it coincides with the monsoon season in the Arabian Sea. All year long, the air is dead still when the sun rises, but it picks up in the early afternoon. By sunset it is blowing so hard that sand stings your face as you walk about; on our expeditions, we needed swim goggles to see well enough to set up our tents. Around midnight the wind drops off again.

*At the point of impact, **a conelike curtain of hot fluid** erupted into the air. The incandescent curtain of molten rock expanded rapidly as more and more of the meteorite made contact with the ground.*

Curtains

Black material and white—the Wabar site offers little else. This dichotomy suggests that a very uniform process created the rocks. The entire impact apparently took place in sand; there is no evidence that it penetrated down to bedrock. In fact, our reconnaissance found no evidence of outcropping rock (bedrock that reaches the surface) anywhere within 30 kilometers.

From the evidence we accumulated

JEFFREY C. WYNN

SECOND-LARGEST CRATER at the Wabar site, Philby "A," has been nearly buried by a creeping seif ("sword," in Arabic) dune. Only its southeastern rim, preserved by a gravelly mix of rock formed during impact, still pokes up above the sand. The 64-meter (210-foot) crater marks the impact site of a five-meter meteorite, one of several pieces of the original Wabar meteoroid (which broke apart in midair). The chunks hit the ground at speeds of up to 25,000 kilometers per hour—20 times as fast as a .45-caliber pistol bullet.

during our expeditions, as well as from the modeling of impacts by H. Jay Melosh and Elisabetta Pierazzo of the University of Arizona, we have pieced together the following sequence of events at Wabar.

The incoming object came from the northwest at a fairly shallow angle. It may have arrived in the late afternoon or early evening, probably during the early spring. Like most other meteoroids, it entered the atmosphere at 11 to 17 kilometers per second (24,600 to 38,000 miles per hour). Because of the oblique angle of its path, the body took longer to pass through the atmosphere than if it had come straight down. Consequently, air resistance had a greater effect on it. This drag force built up as the projectile descended into ever denser air. For most meteoroids, the drag overwhelms the rock strength by eight to 12 kilometers' altitude, and the object explodes in midair. The Wabar impactor, made of iron, held together longer. Nevertheless, it eventually broke up into at least four pieces and slowed to half its initial speed. Calculations suggest a touchdown velocity of between five and seven kilometers per second, about 20 times faster than a speeding .45-caliber pistol bullet.

The general relation among meteorite size, crater size and impact velocity is known from theoretical models, ballistics experiments and observations of nuclear blasts. As a rule of thumb, craters in rock are 20 times as large as the objects that caused them; in sand, which absorbs the impact energy more efficiently, the factor is closer to 12. Therefore, the largest object that hit Wabar was between 8.0 and 9.5 meters in diameter, assuming that the impact velocity was seven or five kilometers per second, respectively. The aggregate mass of the original meteoroid was at least 3,500 tons. Its original kinetic energy amounted to about 100 kilotons of exploding TNT. After the air braking, the largest piece hit with an energy of between nine and 13 kilotons. Although the Hiroshima bomb released a comparable amount of energy, it destroyed a larger area, mainly because it was an airburst rather than an explosion at ground level.

At the point of impact, a conelike curtain of hot fluid—a mixture of the incoming projectile and local sand—erupted into the air. This fluid became the black glass. The incandescent curtain of molten rock expanded rapidly as more and more of the meteorite made contact with the ground. The projectile itself was compressed and flattened during these first few milliseconds. A shock wave swept back through the body; when it reached the rear, small pieces were kicked off—spalled off, in geologic parlance—at gentle speeds. Some of these pieces were engulfed by the curtain, but most escaped and plopped down in the surrounding sand as far as 200 meters away. They are pristine remains of the original meteorite. (Spalling can also throw off pieces of the planet's surface without subjecting them to intense heat and pressure. The famous Martian meteorites, for example, preserved their delicate microstructures despite being blasted into space.)

A shock wave also moved downward, heating and mixing nearby sand. The ratio of iron to sand in the glass pellets suggests that the volume of sand melted was 10 times the size of the meteorite—implying a hemisphere of sand about 27 meters in diameter. Outside this volume, the shock wave, weakened by its progress, did not melt the sand but instead compacted it into "insta-rock": impactite.

The shock wave then caused an eruption of the surface. Some of the impactite was thrown up into the molten glass and was shocked again. In rock samples this mixture appears as thick black paint splattered on the impactite. Other chunks of impactite were completely immersed in glass at temperatures of 10,000 to 20,000 degrees C. When this happened, the sandstone underwent a second transition into bubbly glass.

The largest crater formed in a little over two seconds, the smallest one in only four fifths of a second. At first the craters had a larger, transient shape, but within a few minutes material fell back out of the sky, slumped down the sides of the craters and reduced their volume. The largest transient crater was probably 120 meters in diameter. All the sand that had been there was swept up in a mushroom cloud that rose thousands of meters, perhaps reaching the stratosphere. The evening breeze did not have to be very strong to distribute molten glass 850 meters away.

JOE POLIMENI AND BILL CHASTEEN *Zahid Tractor Corporation and A.M. General Corporation*

EJECTA BLANKET at the edge of the Philby "A" crater consists of three types of debris from the impact: white impactite (a sandstonelike rock formed from compressed sand), black glass (a lavalike rock formed from melted sand) and meteorite fragments (nearly pure iron, with a little nickel). The authors, dressed in special jumpsuits to protect them from the harsh climate, are using magnetometers to search for the meteorite pieces. The tall antenna on the white Hummer vehicle is for Global Positioning System tracking—essential in the middle of the desert, where it is easy to get lost in the protean landscape.

Fading Away

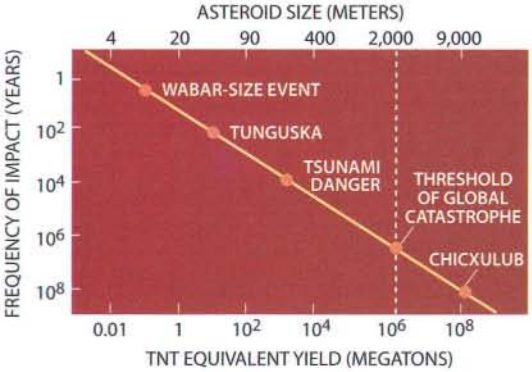
And when did all this take place? That has long been one of the greatest questions about Wabar. The first date assigned to the event, based on fission-track analysis in the early 1970s of glass samples that found their way to the British Museum and the Smithsonian Institution, placed it about 6,400 years

ago. Field evidence, however, hints at a more recent event. The largest crater was 12 meters deep in 1932, eight meters deep in 1961 and nearly filled with sand by 1982. The southeastern rim was only about three meters high during our visits in 1994 and 1995. Dune experts believe it would be impossible to empty a crater once filled.

The Wabar site might have already disappeared if impactite and glass had not anchored the sand. At least two of the craters are underlaid by impactite rocks, which represent the original bowl surface before infilling by sand. We were able to collect several samples of sand beneath this impactite lining for thermoluminescence dating. The results, prepared by John Prescott and Gillian Robertson of the University of Adelaide, suggest that the event took place less than 450 years ago.

The most tantalizing evidence for a recent date is the Nejd meteorites, which were recovered after a fireball passed over Riyadh in either 1863 or 1891, depending on which report you believe. The fireball was said to be headed in the direction of Wabar, and the Nejd meteorites are identical in composition to samples from Wabar. So it is likely that the Wabar calamity happened only 135 years ago. Perhaps the grandfathers of Philby's guides saw the explosion from a long way off.

The date is of more than passing interest. It gives us an idea of how often such events occur. The rate of meteorite hits is fairly straightforward to understand: the bigger they come, the less frequently they fall [*see illustration above*]. The most recently published estimates suggest that something the size of the Wabar impactor strikes the earth about once a decade.



AVERAGE FREQUENCY OF IMPACTS on the earth can be estimated from the amount of scrap material zipping around the solar system and the observed distribution of craters on the moon. A two-kilometer rock, capable of wreaking damage worldwide, falls once every million years on average. (In relating size to explosive energy, this graph assumes a stony asteroid at 20 kilometers per second.)

DAVID MORRISON NASA *Ames Research Center*; LAURIE GRACE

There are similar iron-meteorite craters in Odessa, Tex.; Henbury, Australia; Sikhote-Alin, Siberia; and elsewhere. But 98 percent of Wabar-size events do not leave a crater, even a temporary one. They are caused by stony meteoroids, which lack the structural integrity of metal and break up in the atmosphere. On the one hand, disintegration has the happy consequence of protecting the ground from direct hits. The earth has relatively few craters less than about five kilometers in diameter; it seems that stony asteroids smaller than 100 to 200 meters are blocked by the atmosphere.

On the other hand, this shielding is not as benevolent as it may seem. When objects detonate in the air, they spread their devastation over a wider area. The Tunguska explosion over Siberia in 1908 is thought to have been caused by a stony meteoroid. Although very little of the original object was found on the ground, the airburst leveled 2,200 square kilometers of forest and set much of it on fire. It is only a matter of time before another Hiroshima-size blast from space knocks out a city [see "Collisions with Comets and Asteroids," by Tom Gehrels; SCIENTIFIC AMERICAN, March 1996].

By the standards of known impacts, Wabar and Tunguska are mere dents. Many of the other collision sites around the world, including the Manicouagan ring structure in Quebec, and the Chicxulub site in Mexico's northern Yucatán, are far larger. But such apocalypses happen only every 100 million years on average. The 10-kilometer asteroid that gouged out Chicxulub and snuffed the dinosaurs hit 65 million years ago, and although at least two comparable objects (1627 Ivar and the recently discovered 1998 QS52) are already in earth-crossing orbits, no impact is predicted anytime soon. Wabar-size meteoroids are much more common—and harder for astronomers to spot—than the big monsters. Ironically, until the Wabar expeditions, we knew the least about the most frequent events. The slag and shocked rock in the deserts of Arabia have shown us in remarkable detail what the smaller beasts can do. SA

The Authors

JEFFREY C. WYNN and EUGENE M. SHOEMAKER worked together at the U.S. Geological Survey (USGS) until Shoemaker's death in a car accident in July 1997. Both geoscientists have something of an Indiana Jones reputation. Wynn, based in Reston, Va., has mapped the seafloor using electrical, gravitational, seismic and remote sensing; has analyzed mineral resources on land; and has studied aquifers and archaeological sites around the world. He served as the USGS resident mission chief in Venezuela from 1987 to 1990 and in Saudi Arabia from 1991 to 1995. His car has broken down in the remote deserts of the southwestern U.S., in the western Sahara and in the deep forest in Amazonas, Venezuela; he has come face-to-snout with rattlesnakes, pit vipers and camel spiders. Shoemaker, considered the father of astrogeology, was among the first scientists to recognize the geologic importance of impacts. He founded the Flagstaff, Ariz., facility of the USGS, which trained the Apollo astronauts; searched for earth-orbit-crossing asteroids and comets at Palomar Observatory, north of San Diego; and was a part-time professor at the California Institute of Technology. At the time of his death, he was mapping impact structures in the Australian outback with his wife and scientific partner, Carolyn Shoemaker.

Further Reading

AN ACCOUNT OF EXPLORATION IN THE GREAT SOUTH DESERT OF ARABIA. Harry St. John B. Philby in *Geographical Journal*, Vol. 81, No. 1, pages 1–26; January 1933.

IMPACT CRATERING: A GEOLOGIC PROCESS. H. J. Melosh. Oxford University Press, 1989.

"SECRET" IMPACTS REVEALED. J. Kelly Beatty in *Sky & Telescope*, Vol. 87, No. 2, pages 26–27; February 1994.

HAZARDS DUE TO COMETS AND ASTEROIDS. Edited by Tom Gehrels. University of Arizona Press, 1995.

RAIN OF IRON AND ICE: THE VERY REAL THREAT OF COMET AND ASTEROID BOMBARDMENT. John S. Lewis. Addison-Wesley Publishing, 1996.

Additional information on impact structures can be found at http://bang.lanl.gov/solarsys/eng/tercrate.htm on the World Wide Web.

The Search for Greenland's

Caught on camera, the fireball that streaked across Arctic skies last December appeared to move too fast for anything from this solar system. A monthlong expedition on this island of ice hunted for remains—and answers

TON OF SNOW, melting beneath a blanket of black plastic, was filtered in hopes of finding dust dropped by the 100-metric-ton space rock as it boiled and burst over Greenland last winter.

ANDY RYAN

Mysterious Meteor

by W. Wayt Gibbs, *senior writer*

The astonishing news came via satellite phone, at about 8 P.M., recalled astronomer Lars Lindberg Christensen. He, the four other Danes and the two Greenlanders on the expedition team had just finished a late dinner and were sitting in the communal dome tent, killing time. Time was gnawing back. For seven days, their search for any remnants of the Kangilia meteor had been halted as voices on the other end of the phone repeated variations on the same maddening message: "Stand by.... The helicopter is grounded in Kangerlussuaq by fog.... It's socked in at Paamiut.... It was forced back to Nuuk by the threat of whiteout.... Wait just a few more hours...." Meanwhile the campsite—built on snow that was not even supposed to be on the ice cap this far into Greenland's brief summer—was dissolving into an icy swamp. It was beyond time to move onto the dry, rocky peak of a nunatak and to get on with the hunt.

But now the voice on the phone had good news, incredible news. A television station in Nuuk was reporting that a game warden had found the meteorite. Sailing around the fractal labyrinth of island-dotted coves near Qeqertarsuatsiaat, about 60 kilometers (37 miles) west of camp, the ranger had seen four craters freshly carved from the coastal foothills. Dark rocks lay inside. "It was an intense moment," Christensen recounted the next morning. "Everyone was so excited. We must have burned an hour of satellite time tracking down the guy and arranging for him to guide us to the site." More good news followed: the weather system that had paralyzed the team was breaking up at last. The helicopter would pick them up shortly after dawn to go inspect the craters.

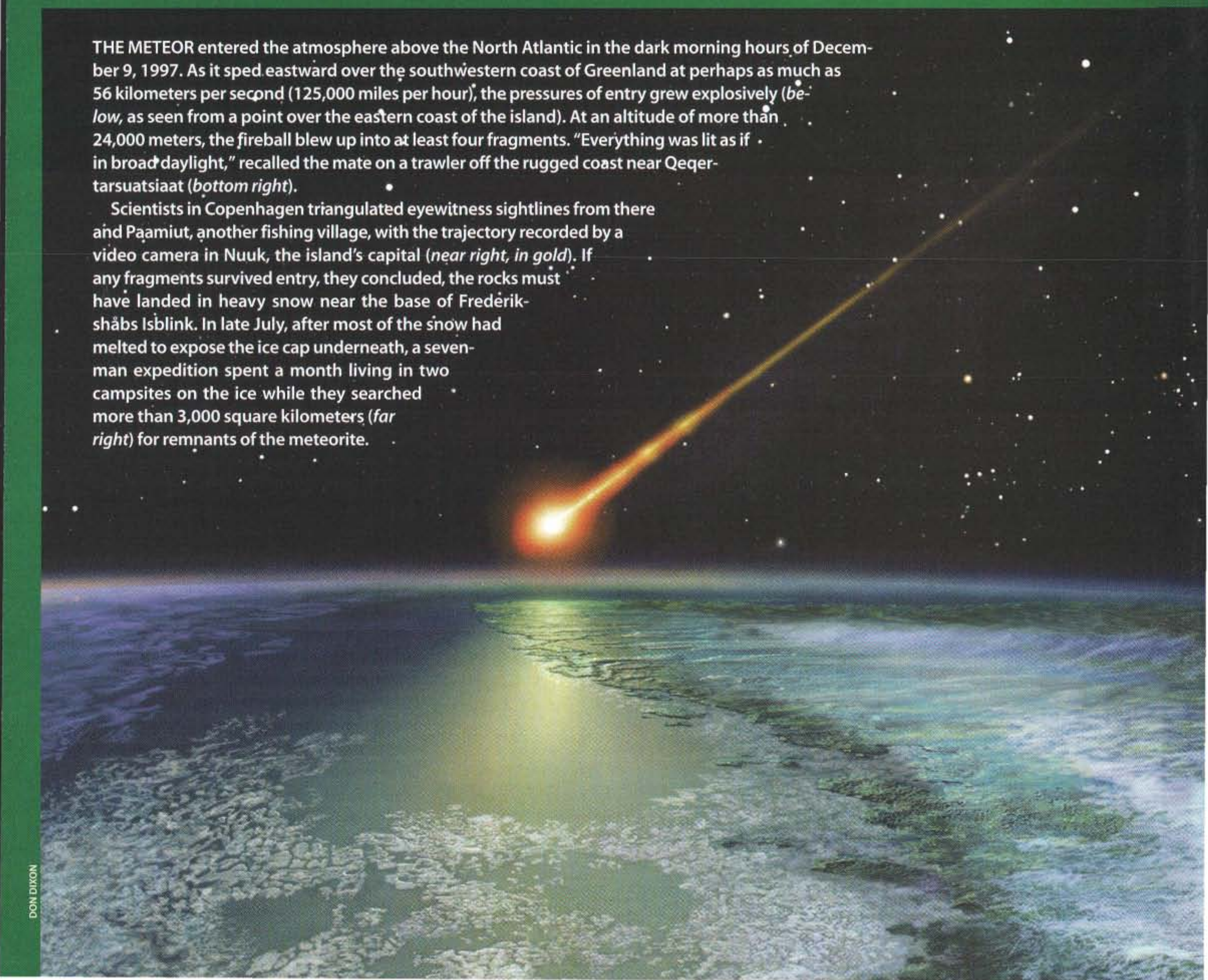
That night Christensen tossed restlessly with anticipation. The fact that they might have spent two weeks trekking over kilometers of ice, climbing in and out of crevasses, melting down snow for its dust, searching every way they could but all in completely the wrong places did not bother him, he said. "Just as long as we—or someone—find something. That's all that matters. Calculations can be wrong. People can make mistakes."

People, he implied but did not say, such as Holger Pedersen and Torben Risbo. An astronomer and a geophysicist at the Niels Bohr Institute for Astronomy, Physics and Geophysics at the University of Copenhagen, they were the scientific brains behind the youthful brawn of this expedition. It was Pedersen and Risbo who had selected the search area, hundreds of square kilometers near the root of Frederikshåbs Isblink, a giant, slow-moving fist of glacial ice that at the campsite is 1,200 meters (almost 4,000 feet) thick. And it was Pedersen and Risbo who, after eight months of detective work, had nearly persuaded themselves and a few other scientists that the fiery meteor that lit up the southwestern coast of Greenland last December 9 had a truly extraordinary origin. Not merely extraterrestrial but extrasolar. Interstellar. The first known ambassador from a whole other star system, perhaps hundreds of light-years away.

That possibility weighed heavily on the 27-year-old Christensen. He knew that if he could bring back the meteorite—or even a pea-size fragment from it—geochemists should be able to prove that unorthodox hypothesis correct. Or, more likely, prove it wrong.

Not that Pedersen seems the kind of researcher who would irresponsibly argue an implausible theory simply to get attention. It is just that, as the 51-year-old scientist earnestly explained to me in his sunny, spartan office in Copenhagen, "it is so extremely unexpected that the first interstellar meteoroid ever detected should be one of this size." Estimates put the mass of the object at

The Kangilia Meteor



THE METEOR entered the atmosphere above the North Atlantic in the dark morning hours of December 9, 1997. As it sped eastward over the southwestern coast of Greenland at perhaps as much as 56 kilometers per second (125,000 miles per hour), the pressures of entry grew explosively (*below*, as seen from a point over the eastern coast of the island). At an altitude of more than 24,000 meters, the fireball blew up into at least four fragments. "Everything was lit as if in broad daylight," recalled the mate on a trawler off the rugged coast near Qeqertarsuatsiaat (*bottom right*).

Scientists in Copenhagen triangulated eyewitness sightlines from there and Paamiut, another fishing village, with the trajectory recorded by a video camera in Nuuk, the island's capital (*near right, in gold*). If any fragments survived entry, they concluded, the rocks must have landed in heavy snow near the base of Frederikshåbs Isblink. In late July, after most of the snow had melted to expose the ice cap underneath, a seven-man expedition spent a month living in two campsites on the ice while they searched more than 3,000 square kilometers (*far right*) for remnants of the meteorite.

DON DIXON

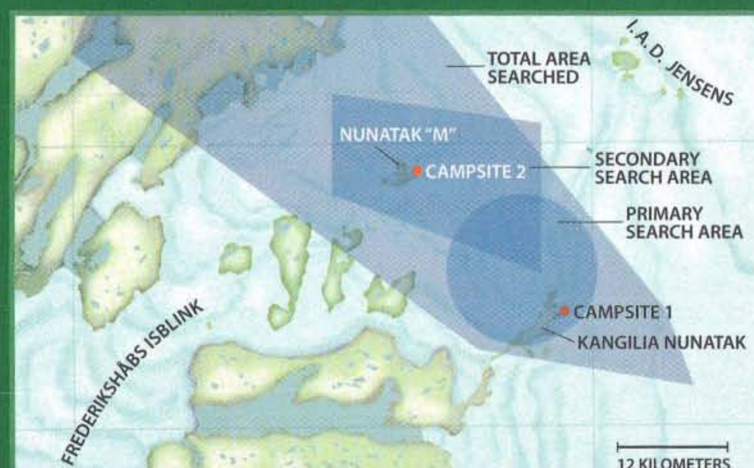
about 100 metric tons before it began to incinerate in the atmosphere. "One would expect very many small interstellar meteoroids to be found before one encounters an object so large," he said, choosing his words with such caution that even his enunciation took on an edge of precision. Yet not one has ever been confirmed. "So we understand perfectly the natural skepticism toward this proposal."

Pedersen also seemed to appreciate the jeopardy in which this tangent away from his specialty—gamma-ray bursts in distant galaxies—could place his career. Zdenek Ceplecha, a highly respected meteor-orbit expert from the Czech Republic, wrote from a conference of meteoriticists that "if I told them (and I have not) that the Greenland meteor was on such a strongly hyperbolic orbit"—that is, one that does not, could not, circle the sun—"they would just start to laugh and not finish until the end of the meeting."

Nevertheless, Pedersen and Risbo had decided to risk ridicule. Or rather they had felt forced to it by the evidence that Pedersen had strewn now across his desk and lap. There was a long strip charting light intensity over time. Several sheets graphing data recorded by American missile defense satellites. A notebook full of scribbled figures, short Danish phrases and hasty, artless diagrams. And on his computer monitor, the object of most of his scientific attention in recent months: the Heilmann video.

Kristian Heilmann had propped a camera inside his home in Nuuk to watch over his snow scooter during the night. In a streak of good fortune, the six-hour tape had caught three seconds of the shooting star's free fall and then, a split second later, one final flare just below the horizon.

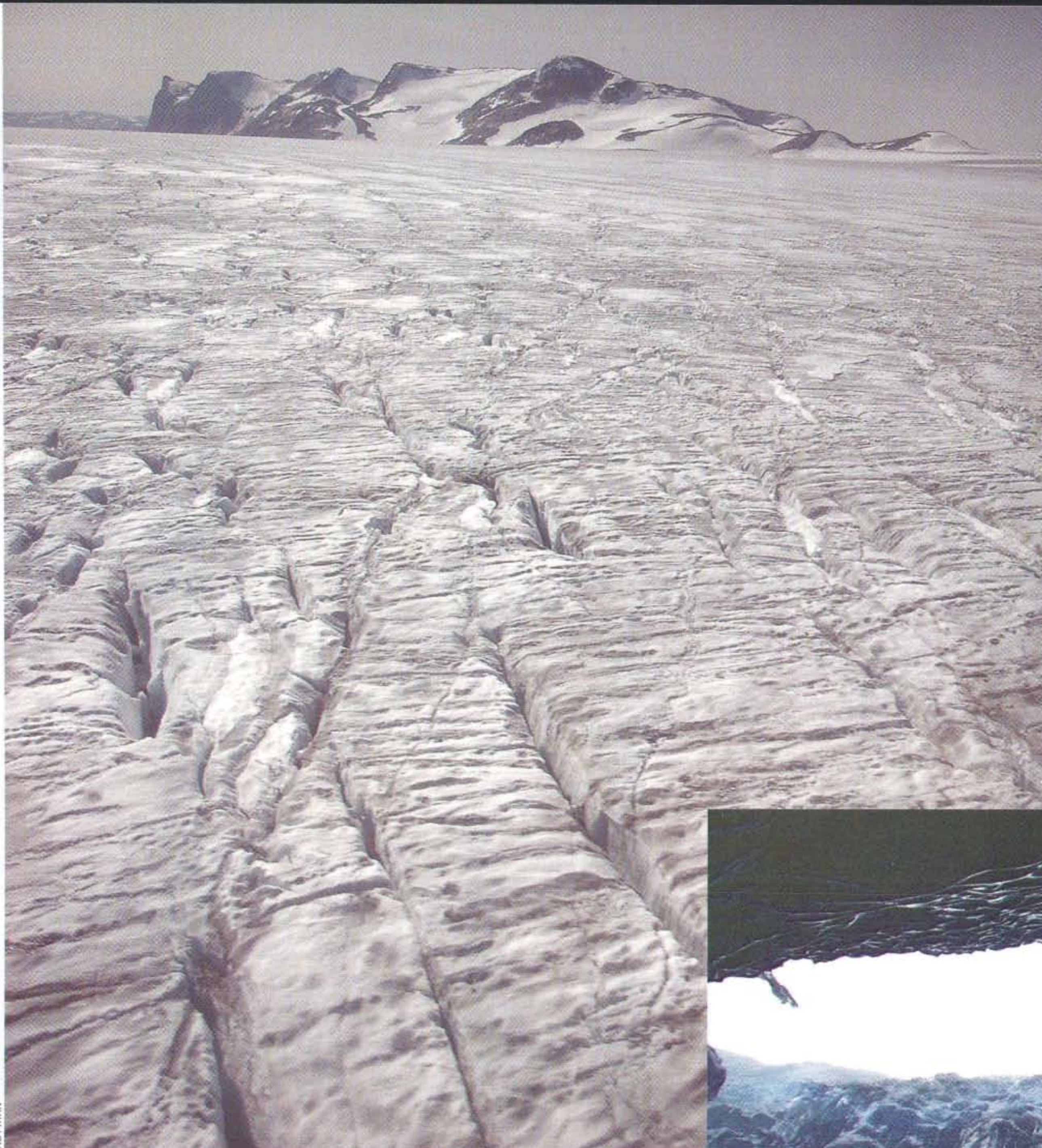
"The value of this video is 100 times that of eyewitnesses," Pedersen said. Cameras have a reliable sense of time. Most observers, on the other hand, reported just what one would expect after the shock of opening fully dilated eyes at 5:15 in the morning to see the



LAURIE GRACE

QEQERTARSUATSIAAT COAST

ANDY RYAN



ANDY RYAN

HANS HENRIK OLSEN

TREACHEROUS ICE in the target zone is fractured with crevasses up to 30 meters wide and 100 meters deep. Despite the dangerous terrain, the team covered about 200 kilometers on foot. Most of the area, however, was searched by helicopter.

energetic equivalent of roughly 64 tons of TNT detonating in midflight. Some described great balls of fire the size of the moon and nearly as bright as the sun. Others recalled trails of flame as wide as two outstretched hands. "*Sunaana!*" cried a three-year-old in Paamiut to her grandmother. "What was that?"

It was a reasonable scientific question, one that Pedersen fancied he might answer. But first he needed to solve two simpler puzzles: In what direction was it traveling and how fast? If he could figure out the course of the fireball as it crossed the camera's field of view, then he would have its speed.

"We knew from the Heilmann video with very high precision that the meteor moved in a plane that tilts up like this at some 47 degrees to the horizon," Pedersen said, excitedly waving a pencil and holding paper sheets edgewise to demonstrate the triangulation of the meteor's path. "But we did not know how the object moved within that plane."

So, for two dark weeks in January, Pedersen had lugged a theodolite, a kind of surveying instrument, from one Greenlandic village to another. Talking to some 30 witnesses, he had filled his notebook with sketches, testimony and, most important, exact measurements of

the angles at which each observer recalled seeing the fireball, taken from the very spot where they saw it.

Meanwhile word of the video was spreading quickly through academic corridors. Ceplecha, intrigued, volunteered his expertise and tools. "We plugged all the constraints into his least-squares program, and that produced the most probable trajectory," Pedersen said. But the same calculations also yielded a most improbable velocity: about 56 kilometers per second. That speed was more than twice the pace at which nearly all meteors enter the uppermost atmosphere; it was especially incredible because the collision was not head-on. This object had practically rear-ended Earth at 125,000 miles per hour.

Tracing the body's path backward, Ceplecha discovered to his dismay that the orbit passed nowhere near the asteroid belt from which most meteors originate. In fact, it was out of the ecliptic altogether, so that even a bizarre slingshot around Mars was out of the question. If the observations and his calculations were correct—which Ceplecha very much doubted—then whatever it was that burned over Greenland must have originated in the deep, deep space well beyond even the Oort cloud of comets, which lies at the fringe of the sun's gravitational well.

Suddenly, the expedition that the Tycho Brahe Planetarium in Copenhagen had already begun planning to hunt for the meteorite took on new importance. If Pedersen and Risbo were correct, there should be nothing to find but dust sprinkled in the snow. The phenomenal pressures of entry would have shattered all but the tiniest fragments. "I think the final destiny of this meteorite was a shower of evaporated material and submillimeter grains," Risbo said. Some of the grains, still buried in up to a meter of *pukak*—year-old snow—might possibly contain proof of an extrasolar origin.

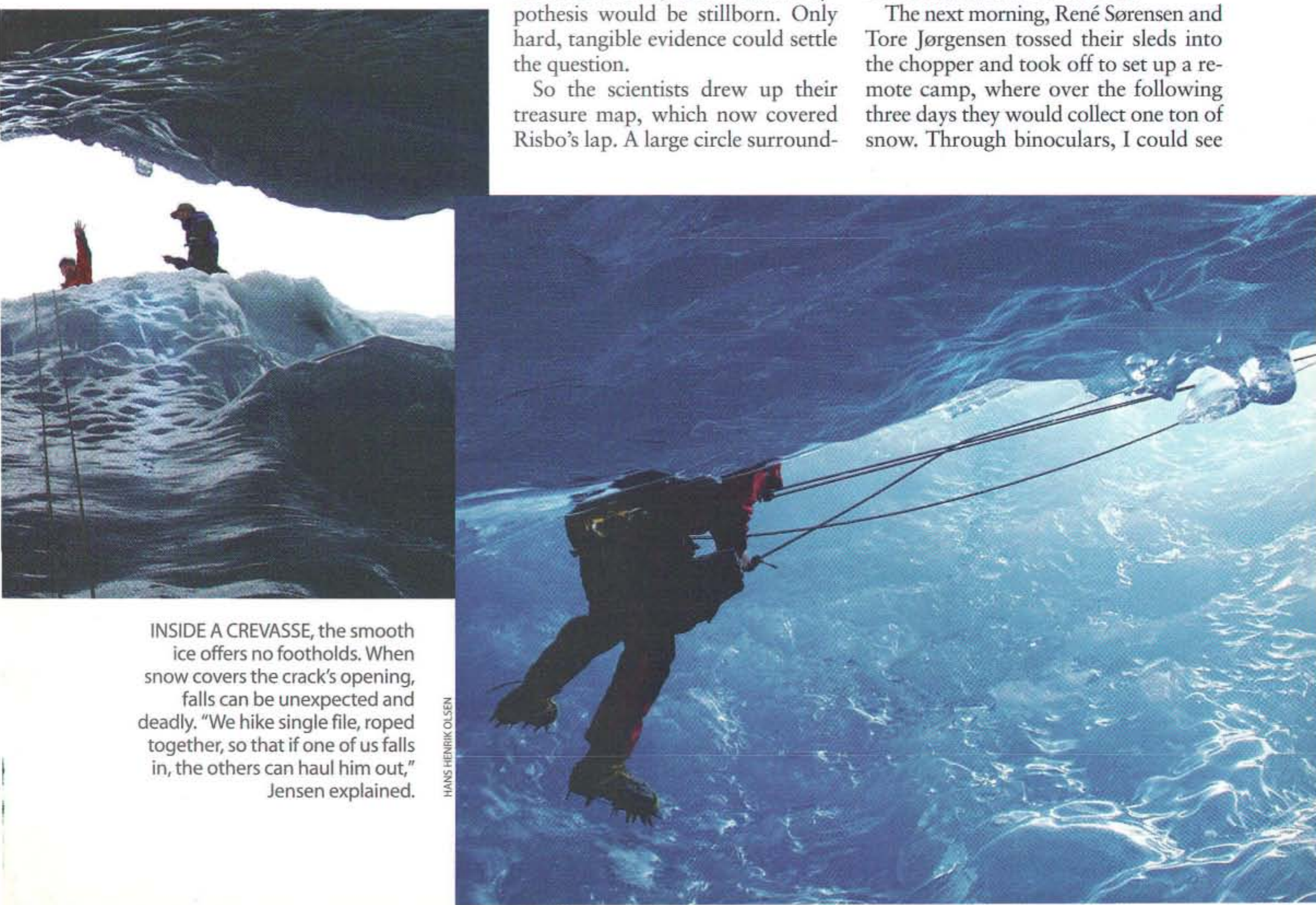
On the other hand, if the hunters returned with a space rock, especially one that fit into one of the known categories of meteorites, the interstellar hypothesis would be stillborn. Only hard, tangible evidence could settle the question.

So the scientists drew up their treasure map, which now covered Risbo's lap. A large circle surrounding Kangilia nunatak marked the primary target area. After a week of searching there, the expedition had found nothing. Now that they finally had a helicopter, however, they had moved camp to the secondary region, an area so remote that the pile of rocks near its center has no official name. Pedersen dubbed it nunatak "M." As in meteor.

*Some described **great balls of fire** the size of the moon and nearly as bright as the sun. Others recalled trails of flame as wide as two outstretched hands.*

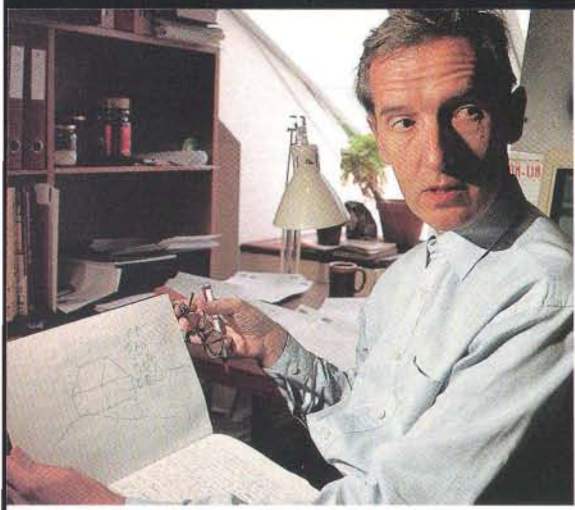
I arrived at nunatak M to find an ambivalent Christensen. He had just returned from the site where the impact craters had been reported. "As soon as I saw them, I could tell they weren't craters," he said. "The slope had just eroded, and some rocks had fallen into the gaps. We're all a bit disappointed." But I could also see the relief in his face as he spied the approaching helicopter. At last, he had transportation. From the air, they would be able to scan more of the ice field in an hour than they had covered in two weeks on foot.

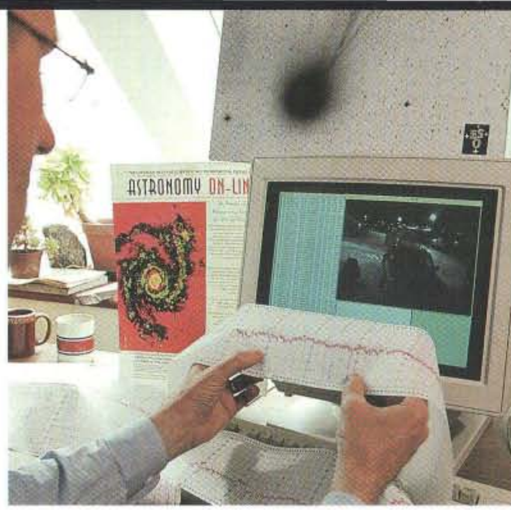
The next morning, René Sørensen and Tore Jørgensen tossed their sleds into the chopper and took off to set up a remote camp, where over the following three days they would collect one ton of snow. Through binoculars, I could see



INSIDE A CREVASSE, the smooth ice offers no footholds. When snow covers the crack's opening, falls can be unexpected and deadly. "We hike single file, roped together, so that if one of us falls in, the others can haul him out," Jensen explained.

HANS HENRIK OLSEN





ANDY RYAN

ASTRONOMER Holger Pedersen went door to door collecting testimony from Greenlanders who witnessed the falling star. By correlating their accounts (*left*), the video recorded in Nuuk (*right*) and data from U.S. Space Command satellites, he arrived at the extraordinary conclusion that the object was the first meteor ever observed to arrive from interstellar space.

them unrolling the black plastic, shoveling the *pukak*, stiff as snowcone snow, onto it. Later they would filter the nearly pure meltwater for whatever trace solids had been trapped in the snow since it started accumulating last August.

"The first week, we tried examining the particles under a microscope, to see whether any looked promising," Christensen said. "It was storming all the time, so we worked inside the tent. But the minute we opened up the samples, they got contaminated"—by soot from their kerosene burners, fur from their reindeer skins and hair from their heads. All samples since have been sealed for later analysis in Copenhagen.

That study will take months to complete and may be more scientifically challenging than fieldwork in Greenland. But it would not offer half the thrill. My senses reeled at the vista as I looked out past the two diggers in the distance, mere fleas on a beach of sugar; out past the I.A.D. Jensens nunataks, five huge mountains submerged like a herd of whales below an ocean of white, only their gray dorsal fins tentatively slicing through the surface; out perhaps 90 kilometers to a horizon defined solely by the curvature of the planet.

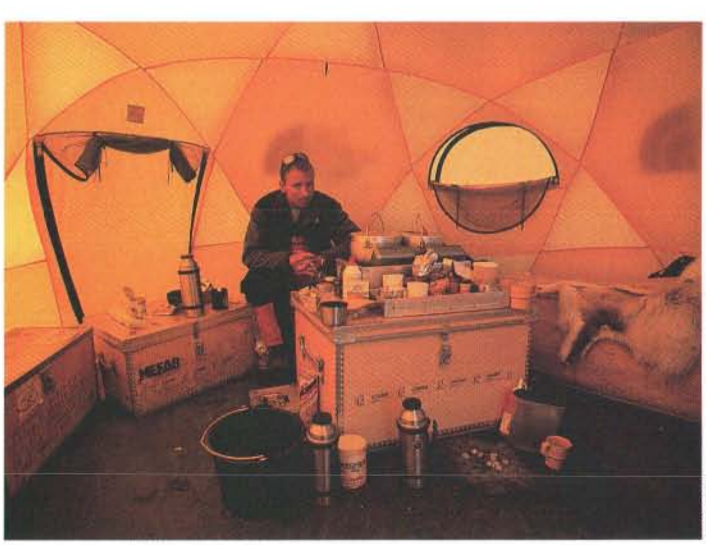
Claus Kongsgaard Jensen, a pilot with the Danish Air Force, stood next to me. Like all the others on the team, he was an unpaid volunteer who had traded a year's worth of vacation time to be here. He followed my northeasterly gaze with eyes the same eerie shade of green as the glacier-fed fjords near Paamiut. "You can go that way for about 1,000 kilometers without meeting anything other than ice," he said, smiling. After the I.A.D. Jensens, the ice thickens to 3,300 meters, swallowing the mountains whole.

By 9 A.M. the men had finished their second breakfast. Six of them jumped into the helicopter. The aerial hunt had begun. The aircraft took off toward the southeast, where Ceplecha, Risbo and Pedersen had figured the largest fragments probably hit, if they landed at all. The plan was to cover several hundred square kilometers in about two hours, using a lattice search pattern: forward 10 kilometers, right one, back 10, left one, repeat, repeat, repeat. "Like mowing the lawn," Jensen said.

From an altitude of 30 meters six eyes intently scanned a 900-meter swath of ice on the left side of the chopper. Six more eyes scanned right. Even at 50 knots, Christensen said, "if something was out there, and it wasn't buried in the snow, then we would see it. I'm very confident of that."

After they had disappeared from view, I sat down on nunatak M, a jumble of dark granite, rusty iron ore and crystalline white quartz crushed to shards by fracturing cold, crawling ice and three billion years of age. Everything suddenly came to a standstill, as if someone had just switched off time.

Coming so soon after the deafening chop of the Bell 212's massive blades, the silence was a little frightening—too perfect. There were so many sounds missing: traffic, jets, human movement, birds, insects, wind. This is what it must be like to be deaf, I thought.



ANDY RYAN

WILD GOOSE CHASE took the team off the ice to investigate a report of impact craters spotted on the coast. But the pits (*white spots, above*) turned out to be natural erosion features. That misdirection, a week of paralysis imposed by foggy weather and a fruitless aerial hunt took its toll on the spirits of team astronomer Lars Lindberg Christensen (*right*).

I heard the chopper before I saw it, from at least 20 kilometers away. When it landed, Christensen emerged holding a bag full of gooey sand, black as obsidian. "We found nothing," he said glumly. He held up the bag. "I saw this stuff sitting above the surface of a melt pond. But it's only a cryoconite." Eddies and pools in the meltwater brooks that run through and beneath the snow cover collect dust and algae swept up by the streams. Over time, he explained, the black deposits grow into cryoconites, piles up to a meter across.

"There may be dust from the meteorite in here," he said. And yet the thought did not seem to cheer him. At dinner the previous night I had asked Christensen what he expected to find. "I am totally objective," he claimed. "I have no expectations."

But Christensen now looked mighty sullen for a man with no expectations. "It's exhausting. We gave it all the skill and attention we have," he said, a bit defensively, as if his boss back at the planetarium were listening.

That afternoon the team went up for another 90-minute survey. It came down empty-handed. On the third and final trip of the day, Jensen called twice for the pilot to descend. But each time, hopes for a rock dissolved into black piles of cryoconite goo.

By the end of the day Christensen was ready to admit that he was a human being with an opinion after all. "It's a black-or-white kind of expedition," he said. "Either we find the meteorite or we don't. Right now I see only the black."

The next day, at Sandia National Laboratory in New Mexico, Ceplecha met with remote-sensing specialist Richard E. Spalding. They were trying to reconcile the Heilmann video with the data from two U.S. Space Command satellites that had also recorded the meteor's explosive breakup.

Pedersen had been claiming that the video did not distort time, and Sandia technicians had just arrived at the same conclusion. "Two flashes in the video are strongly, unmistakably correlated with two very bright peaks in the satellite record," Spalding said.

ANDY RYAN

VASTNESS of the ice field made for slim odds of finding the meteorite. "We had to go look anyway," Christensen said. "There was no way out of it."

Spalding called Edward Tagliaferri, a consultant often tapped by the Pentagon to analyze classified satellite observations—especially odd ones. "My reading of the satellite data is that the flashes indicate a velocity of 46 to 52 kilometers per second," Tagliaferri said. "It could not be half that. But it is possible that I'm being fooled somehow." Perhaps, he suggested, two large chunks, separated by several kilometers, blew up within milliseconds of each other.

Spalding offered another, equally speculative, possibility. What if the meteor's ionized trail had connected two charged layers of the atmosphere like a wire, he wondered. "A plasma of accelerated ions might then have rushed down, passed the body and kept on screaming forward a ways until they dumped their energy all at once in a giant flash. Like a lightning bolt, but made of ions."

The video and satellite cameras might have seen a plasma bolt zipping ahead of the much slower (and more conventional) meteor, Spalding said. It's an unorthodox idea, but Ceplecha said that he, for one, is more inclined to believe in plasma bolts than interstellar meteors.

A month later, in early September, Pedersen began to suspect that perhaps neither of these outlandish explanations was necessary, because perhaps they had been wrong about the speed all along. He had just received a new digital copy of the video. It was much clearer than the first, so clear in fact that he could precisely match frames in the tape to blips in the satellite record. With this new calibration, Pedersen painstakingly recalculated the fireball's path and arrived at a much different answer than Ceplecha had: a speed of 29 kilometers per second, not 56; an origin safely within the solar system, not beyond it.

Although his new model sounded more reasonable, it still produced one confusing conclusion, Pedersen said. If the rock was moving at a normal speed, then it should have been visible very early in its entry. "Almost all meteors begin shining between 110 and 90 kilometers altitude," he pointed out. Yet this fireball apparently didn't heat up until it reached 70 kilometers. That is hard to explain.

Presented with such unsatisfying alternatives, most meteoriticists will probably throw up their hands and quickly dismiss the Kangilia meteorite as exceptional in any way save its size. Unless, of course, the expedition brought back solid, physical evidence to the contrary.

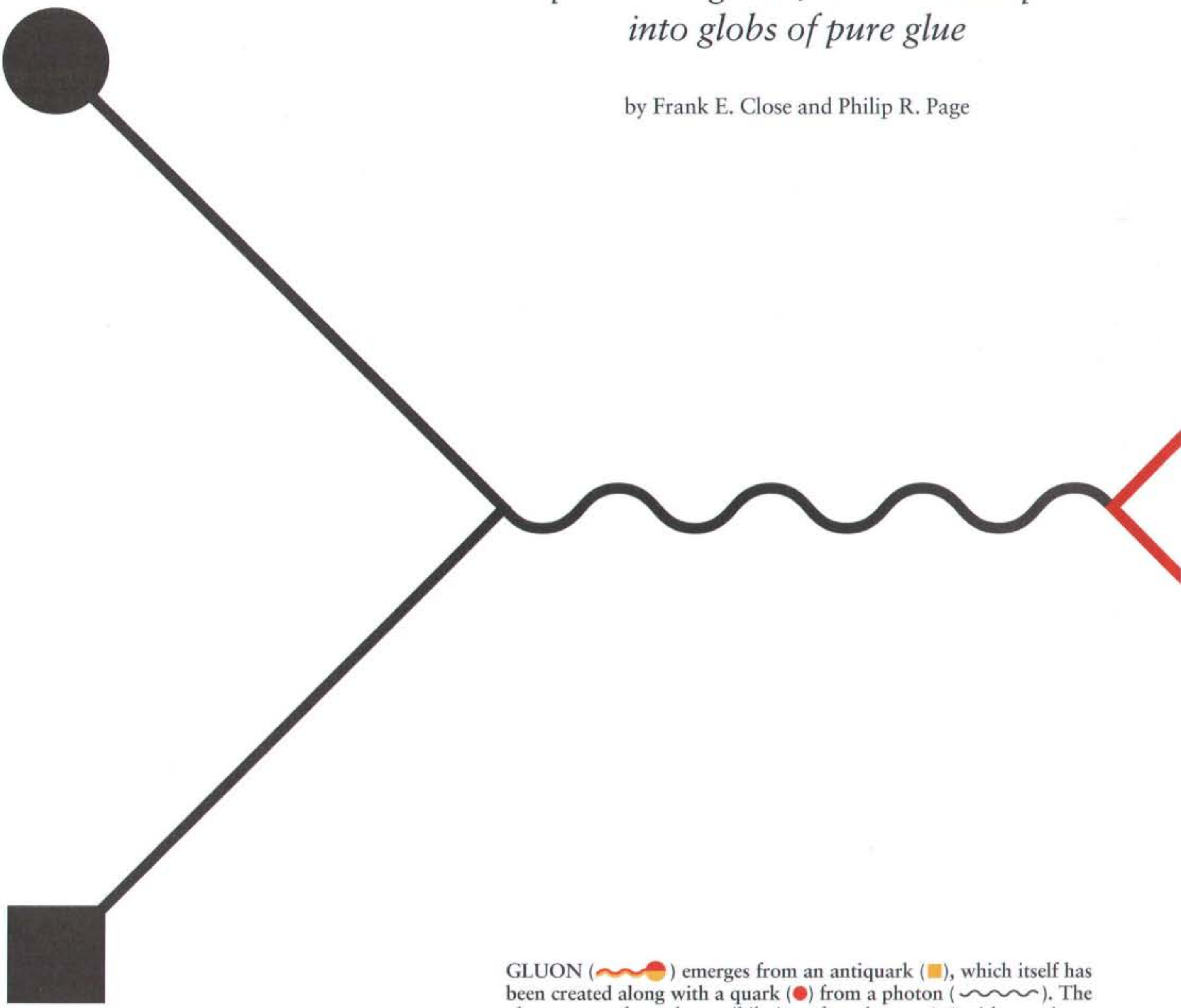
On August 16 the expedition ended as it began, thwarted by the capricious Greenlandic weather. By the time the fog rolled in and forced them off the ice, the team members had walked some 200 kilometers in rain, wind and snow over the treacherous ice cap. They had overflown more than 1,000 kilometers, searching in 25 days an area the size of Rhode Island for a rock the size of a golf ball.

Perhaps they had captured some telltale grains of falling stardust in their snow samples. But it seems more likely that the true nature of the Kangilia meteor will escape science's grasp. Snow started falling over the ice on the day the team left. Soon any fragments lying in melt ponds, at the bottom of blue crevasses or among the stony litter of glacial moraines will be buried, and Greenland will add another secret to the many trapped under its cap. SA

Glueballs

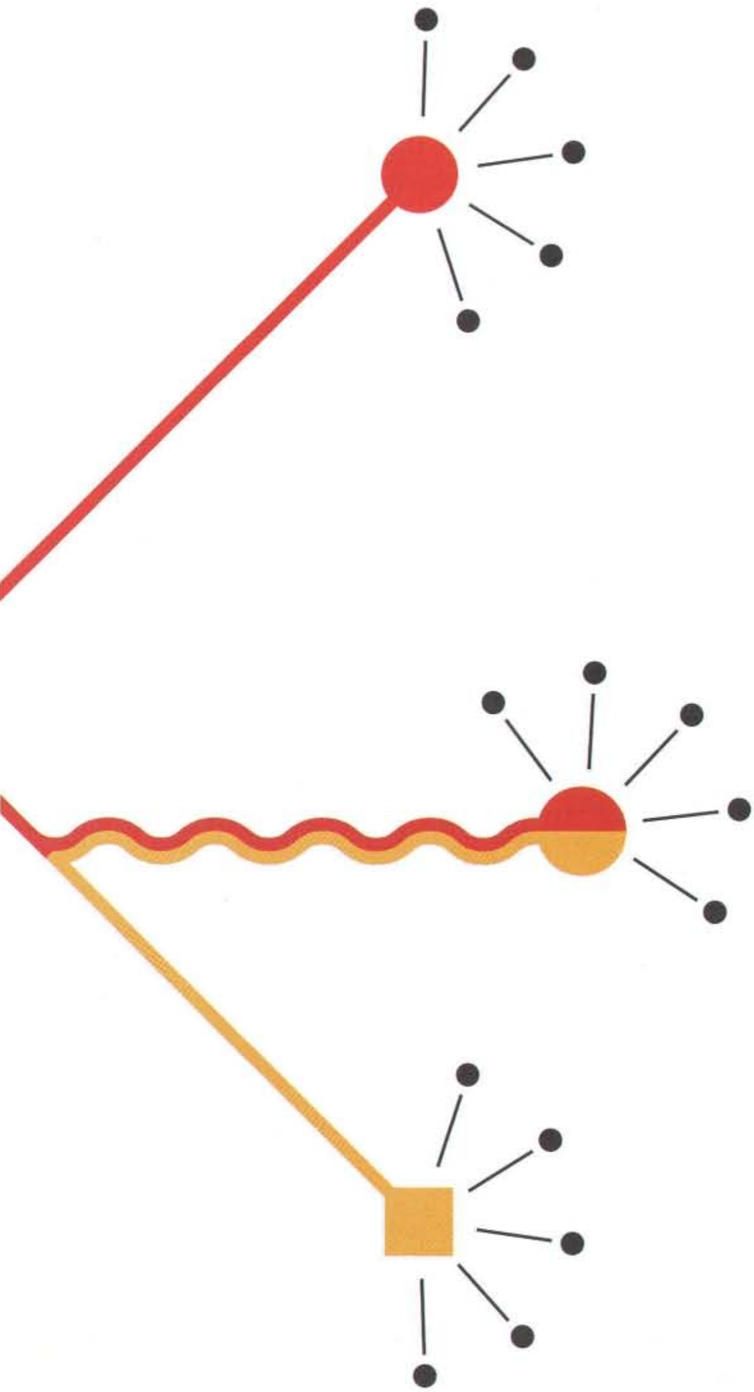
Gluons, which hold protons together, can also clump into globs of pure glue

by Frank E. Close and Philip R. Page



GLUON ( ) emerges from an antiquark ( ), which itself has been created along with a quark ( ) from a photon ( ). The photon came from the annihilation of an electron ( ) with a positron ( ). A quark, an antiquark and a gluon carry "color" charges that allow them to interact via the strong nuclear force. A color and its anticolor (or opposite) charge are depicted here by the same hue; the convention leads to every colored line being continuous. Because nature abhors color, the quark, the antiquark and the gluon combine with other particles to yield showers, or "jets" ( ), of colorless objects that are observed in experiments. The electron, the positron and the photon do not have color charge and are not subject to the strong force. Time flows from left to right in this Feynman diagram; such drawings were invented by Richard Feynman as an aid to calculation.

BRYAN CHRISTIE



There are no atoms of light. That is, photons do not attach to other photons, forming composite entities. But gluons, the particles that bind quarks—the basic units of matter—into objects such as protons, may indeed stick just to one another. Physicists call the resulting glob a glueball.

A glueball is thought to have a radius of 0.5×10^{-15} meter, less than that of a proton, and live for less time than light takes to cross a hydrogen atom. Ephemeral though these particles may seem, in the past year many physicists have become convinced that glueballs are showing up in experiments.

Finding a glueball would be a thrilling culmination to a colorful story. By the 1960s scientists had realized that some of the particles they observed in experiments are compounds of smaller objects called quarks. Along with electric charge, quarks have other, odder attributes. One kind—called flavor—comes in six varieties whimsically named up, down, charm, strange, top and bottom. Another kind of property is called "color" charge. An up quark, for instance, comes in three basic colors: red, yellow or blue.

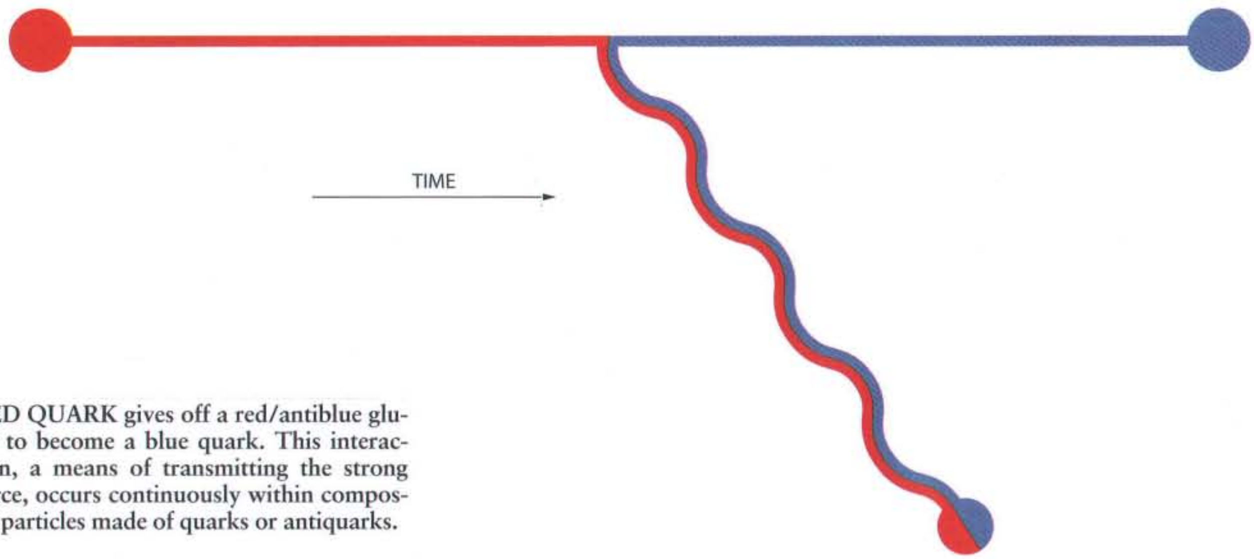
A blue quark will bind with a red quark and a yellow quark, forming a "white" object that has no color charge. The result may be a proton, a neutron or any one of innumerable three-quark composites called baryons. (Physicists call these charges "color" because all three of them add up to zero, just as the three primary colors add up to white.) The attraction between different colors is in fact the so-called strong nuclear force, which binds quarks into these larger, stable objects.

Quarks may also attach to antiquarks, particles that have opposite charges to the quarks. An antiquark comes in anticolors—antired, antiyellow or antiblue. (An anticolor is mathematically denoted by negative color. Antired, for instance, is a "minus red.") Opposite colors attract. A red quark, for instance, will bind with an antired antiquark to form a white object called a meson. The most common meson is a pion, often observed in nuclear reactions.

Electric and color charges are analogous in other ways. The theory of electromagnetism describes the attraction between opposite electric charges. In the 1940s physicists merged electromagnetism with relativity and quantum theory, creating a branch of fundamental physics termed quantum electrodynamics, or QED. This theory—the most successful known to physics—holds that the electromagnetic force is transmitted by massless objects called photons. These quanta of light banish the classical idea of action at a distance: they bounce, say, between an electron and its antiparticle, the positron, in such a manner as to draw the two together.

The equivalent theory of color charges, which communicate via the strong force, is called quantum chromodynamics, or QCD. Gluons, the massless quanta of the strong force, transmit this interaction.

Gluons are fundamentally different from photons. Even though photons allow particles of opposite electric charge to attract, they themselves do not have charge. So one photon cannot push or pull on another. Gluons, however, are themselves colored. A red quark, for instance, can turn into a blue quark by radiating a red/antiblue gluon. Basically, a gluon can attract another gluon.



RED QUARK gives off a red/antiblue gluon to become a blue quark. This interaction, a means of transmitting the strong force, occurs continuously within composite particles made of quarks or antiquarks.

BRYAN CHRISTIE

This behavior makes gluons peculiar compared with their electromagnetic cousins. Photons surround an electron uniformly, forming a shell with spherical symmetry. Moreover, the density of photons falls off with distance, so that the attraction between an electron and positron diminishes with the inverse square of the distance between them. Gluons are not so uniformly distributed. Instead they clump together into a tube linking a quark and an antiquark. The color originating in the quark can be thought to "flow" through the gluon tube to the antiquark, where it becomes absorbed.

A Tube of Glue

When the tube is stretched out, it pulls back with a constant force, whatever its length. An infinite amount of energy would be required to stretch the tube to infinite length—that is, to pull apart a quark and an antiquark. As a result, free quarks never occur: nature abhors color. But a gluon tube holding a meson together may break into two pieces. The colored ends of the new tubes correspond to quarks (or antiquarks), so that one meson breaks into two [*see illustration on opposite page*].

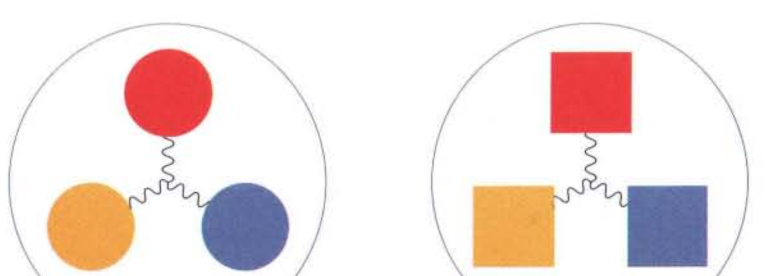
Gluons acting in this way, at least in theory, do fit the observations. But are they real? The clearest evidence for gluons comes from the annihilation of an electron with a positron. The energy released when these particles combine is often reborn as a quark and an antiquark that initially fly in opposite directions. The tube between them breaks up into a shower of mesons and baryons. So an experimenter sees jets of the composite particles emerging back to back.

Sometimes, however, the quark or antiquark may emit a gluon, giving rise to a total of three jets [*see illustration on preceding page*]. Researchers at DESY (German Electron Synchrotron) in Hamburg first observed these triads in 1979. Happily, they had just the properties that QCD theorists had predicted. Such experiments have also confirmed that gluons can emit other gluons.

Other investigations at DESY, in which high-energy electrons penetrate protons to survey their internal structure, have discerned the presence of gluons. The particles reveal themselves in other ways as well. For instance, mesons made of heavy quarks or antiquarks have the masses expected if the quantum of color force has an intrinsic angular momentum of 1—as a gluon does.

But gluons may in fact manifest themselves directly. In 1972 Harald Fritzsch and Murray Gell-Mann, both then at the California Institute of Technology, predicted that two or more gluons—say, a red/antiblue and a blue/antired—can combine into a strongly bound, neutral-colored particle of pure glue. This hypothetical object is called a glueball. The theorists also argued that a gluon can bind with a meson to form a hybrid. For instance, a red quark and an antiblue antiquark can conceivably bind with a blue/antired gluon, forming a white combination.

It was an elegant idea. Quantum chromodynamics is, however, a rather messy theory: the peculiar, sticky character of the strong force makes it impossible for physicists to carry out exact calculations. Almost everything known about color and glue comes not from direct calculation but from massive computer simulations known as lattice QCD [see "Quarks by Computer," by Donald H. Weingarten; SCIENTIFIC AMERICAN, February 1996]. This approach regards space-time as a grid, called a lattice. A quark might sit at one point of the lat-



BRYAN CHRISTIE

BARYON (*left*) is overall neutral in color, or "white." The most common baryon is the proton or the neutron. The quarks within a proton interact with one another by exchanging gluons and perpetually altering their color. An antibaryon (*right*), such as an antiproton, contains an antired, an antiblue and an antiyellow antiquark.

tice, being connected by a line with an antiquark at another point. In such computations, an entity made of pure glue does turn up: a gluon line that closes in on itself like a snake eating its own tail. This loop is a crude representation of a glueball.

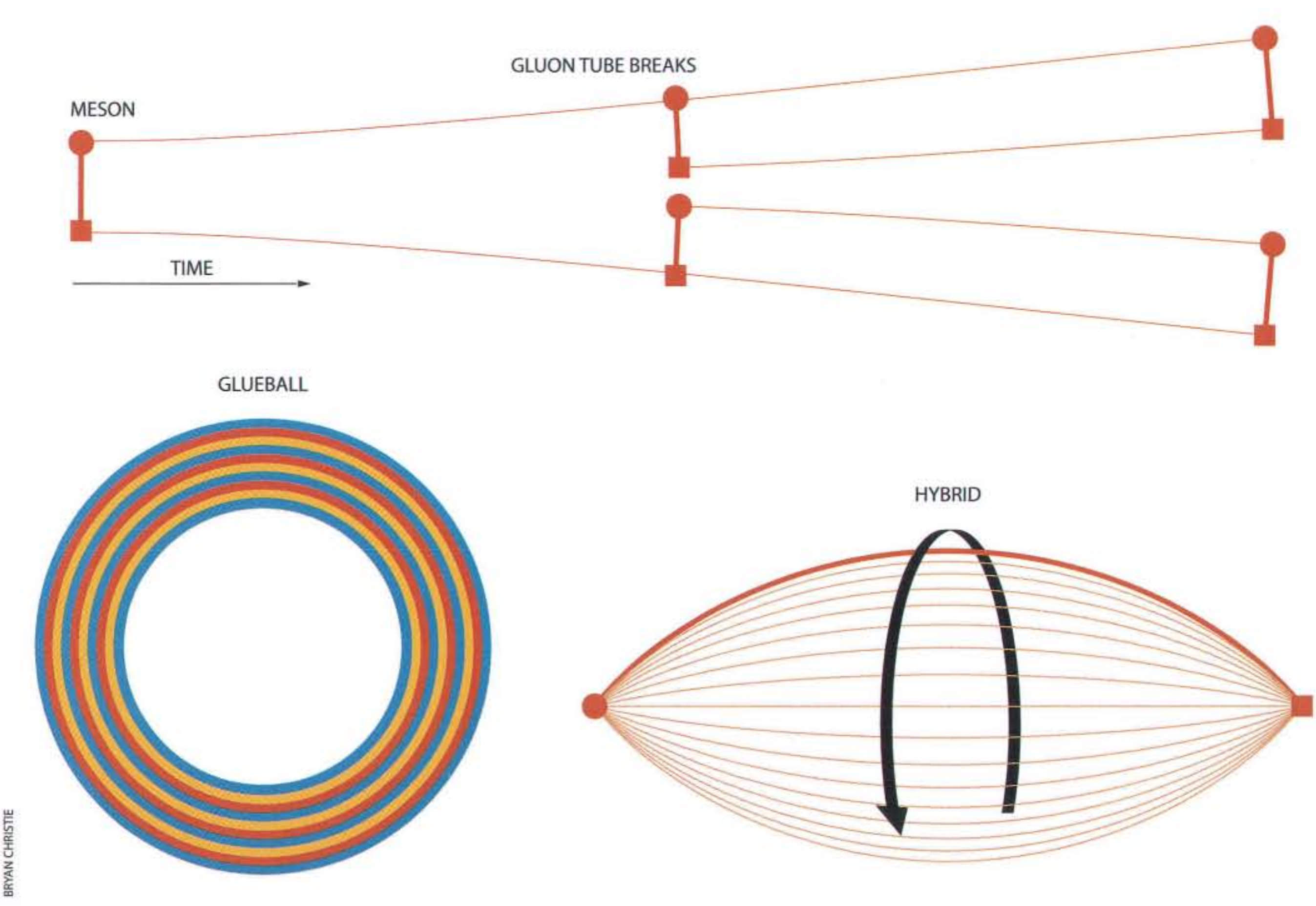
The lightest glueball allowed by theory corresponds to a circular tube of glue and has no angular momentum. According to quantum mechanics, an object without angular momentum must have the symmetry of a sphere. And in fact the glue ring can be oriented in any direction, so that the wave function describing it ultimately corresponds to a spherical shell. Glueballs of other, elongated shapes have nonspherical wave functions. They therefore have angular momenta and larger masses.

Sometimes the stringlike tube of glue connecting a quark-antiquark pair is slack, so that it does not lie along the axis connecting the pair but hangs off it. It may then spin about this axis like a jump rope. Such configurations can be thought of as containing an extra gluon, corresponding to the aforementioned hybrids. Because the rotation requires considerable energy, the total internal energy or mass of a hybrid will be larger than that of a meson alone.

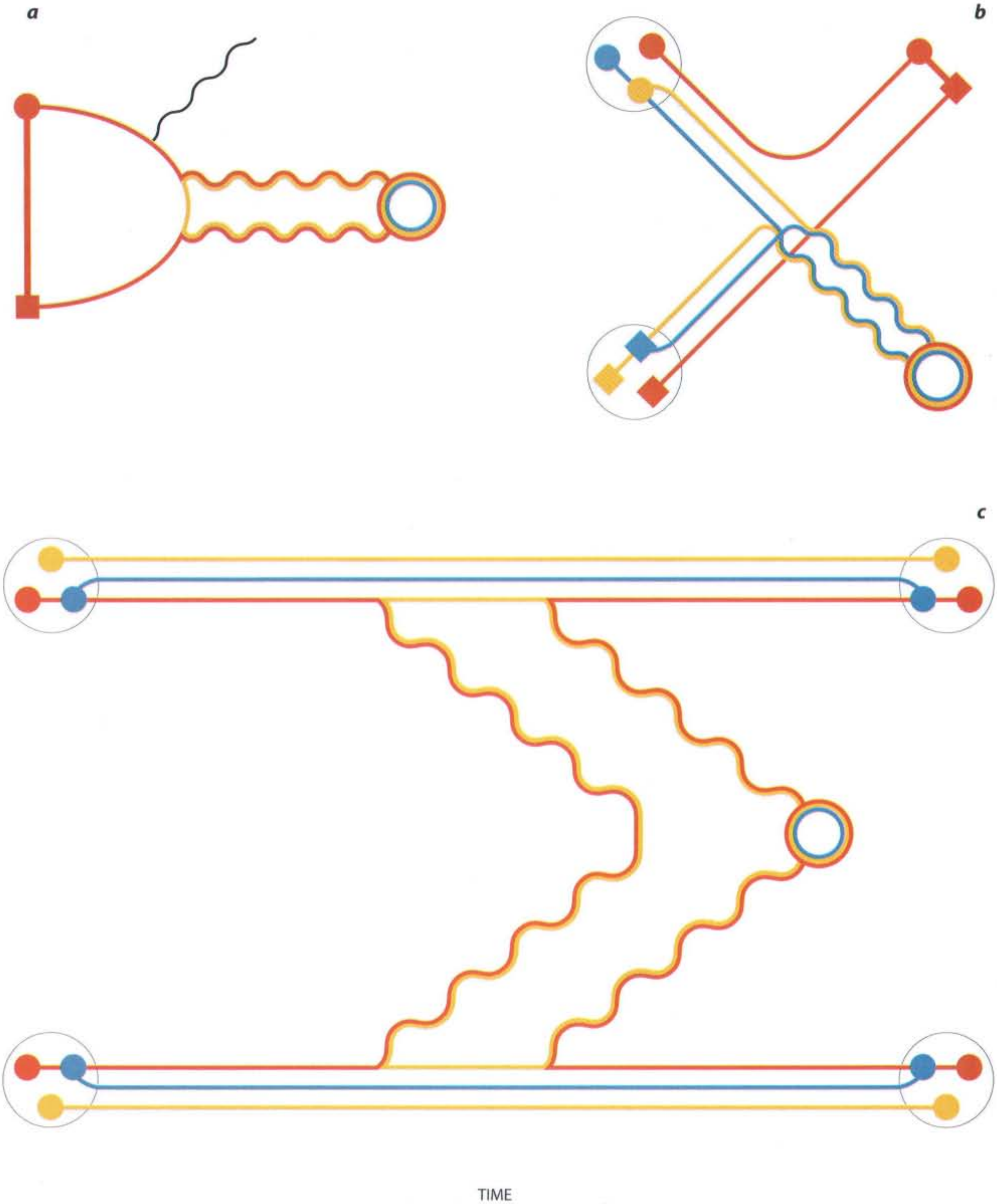
An ordinary meson (one not linked to an extra gluon) can also rotate, but it does so about the axis perpendicular to the line joining the quark and the antiquark that this particle contains. The mesons of lowest energy have no such rotation and are described as S-wave. Typically an S-wave meson (such as the commonplace pion) will decay into two S-wave mesons. A hybrid rarely decays in this manner, however, because the angular momentum it possesses by virtue of its peculiar rotation must be conserved. Instead hybrids are predicted to decay into one S-wave meson and another short-lived meson that does have some internal angular momentum. The latter then breaks up into two S-wave mesons. In sum, the hybrid should announce its presence by yielding at least three S-wave mesons.

For many years, the two of us could only paraphrase Mark Twain: "Everybody talks about hybrids, but nobody does anything about them." Experimenters find it difficult to detect all three mesons emitted in a decay and thus rarely have sought to observe them.

In December 1994 we finished a calculation of all possible decays of hybrids, using a model in which the glue tube is made up of a chain of balls representing gluons. The model yields results similar to those of lattice calculations, but it can



MESON, GLUEBALL AND HYBRID are three composites involving a tube of gluons. A commonly observed type of particle, a meson consists of a quark and an antiquark connected by a gluon tube. Color can be thought to originate in the quark and "flow" to the antiquark. If the tube should break at some point, the colored ends serve as a quark and an antiquark, giving rise to two mesons. A glueball appears in calculations as a ring of pure glue. Because the colors of the gluons cancel one another, in theory a glueball should be quite stable. A hybrid, another hypothetical particle, is essentially a meson in which the gluon tube rotates about a line connecting the quark and the antiquark. It can be thought of as a meson allied with an extra gluon.



BRYAN CHRISTIE

PROCESSES RICH IN GLUE are promising places to seek glueballs. A ψ meson may decay by giving off two gluons, which join into a glueball (*a*). Or a proton might crash into an antiproton, so that the quarks of the former interact with the antiquarks of the latter, emitting two gluons and a meson (*b*). The gluons may combine into a glueball. Moreover, when two protons collide head-on, their constituent quarks communicate by interchanging gluons (*c*). Two of these gluons could merge into a glueball. The glueball may itself decay in diverse ways, yielding two pions or other mesons (***not shown***).

illuminate more situations. (Lattice calculations cannot as yet describe the decays of hybrids, but they should be able to do so quite soon.) To our excitement, we realized that a hybrid might have been discovered earlier that year.

In the summer of 1994 experimenters based at the Institute for High Energy Physics in Protvino, Russia, had reported an object called $\pi(1800)$ emerging from collisions of pions with protons. (The number in parentheses denotes the mass in million electron volts, or MeV.) This particle has the quantum characteristics and pattern of decay expected for a hybrid.

Exotics

Unfortunately, the identity of $\pi(1800)$ still remains uncertain. But recent evidence has emerged for the existence of special hybrids called exotics. Mesons are characterized by several quantum numbers: their internal angular momentum (J), their appearance when reflected through their center (parity, or P) and their identity when particles are interchanged with antiparticles (charge conjugation, or C). Exotics have combinations of these three quantum numbers that are forbidden for mesons. The simplest exotic has $J = 1$, $P = -1$ and $C = +1$. (The commonplace meson with $J = 1$ has $P = -1$ and $C = -1$.)

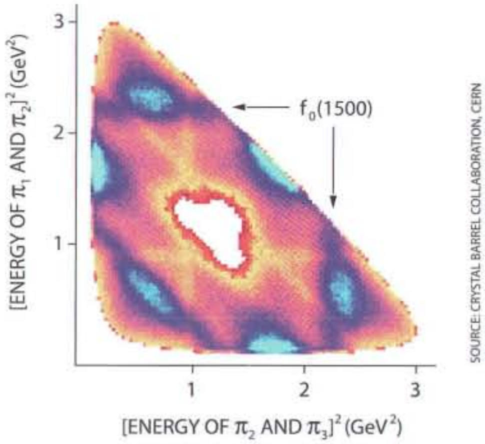
In 1997 experimenters at Brookhaven National Laboratory, also working with pion beams and proton targets, reported possible evidence for this exotic. Soon after, the Crystal Barrel collaboration at CERN, the European laboratory for particle physics near Geneva, also announced signs of the same exotic. These objects were, however, not observed by their breakup into three S-wave mesons. Instead the experimenters searched for two S-wave mesons, the η (eta) and the π (pion). (These are not the most abundant products of an exotic's decay but are simply those that can be more readily detected.) Most recently, the Brookhaven group has reported evidence for an exotic of a different mass. Whether or not these objects are actually exotics remains a matter of much argumentation.

Glueballs are proving even harder to find, because most experiments check for final products made up of quarks and antiquarks. In 1993 a collaboration based at the University of Edinburgh predicted a mass of 1,550 MeV for glueballs using lattice theory. Soon after, Donald Weingarten of the IBM Thomas J. Watson Research Center and his co-workers calculated a mass of 1,740 MeV. Later they predicted that glueballs should live for 10^{-24} second. This minuscule amount of time is actually long enough for glueballs to be detected.

In conferences of the time, one of us (Close) emphasized that intensive experimental effort should be devoted to searching for glueballs of mass in the range 1,500 to 1,800 MeV—without knowledge that the Crystal Barrel collaboration was starting to uncover a short-lived object known as $f_0(1500)$. Its angular momentum (denoted by the subscript) is 0, whereas its mass is within the predicted realm. This object could be the lightest glueball.

Or maybe not. For more than a decade, experimenters have had evidence of an object called $f_J(1710)$, whose angular momentum J is still undetermined. If J turns out to be 0, this object will be a rival candidate for a glueball.

As if this ambiguity were not confusing enough, the situation is worsened by quantum mechanics, which dictates that two objects of identical quantum numbers and similar masses can mix with each other. Calculations predict two mesons of the same quantum numbers as the glueball in the same region of mass. The glueball should combine with these mesons to produce three final objects that are partially glue. Indeed, the $f_0(1500)$ and $f_J(1710)$ do seem to share features that would incriminate them as being about half glueball. Unfortunately, the evidence remains inconclusive.



DALITZ PLOT of a collision between a proton and an antiproton may encode evidence of a glueball. The reaction yields three pions, two of which may have come from the breakup of a glueball. Each axis plots the square of the combined energy of two of the three pions taken together. A region of enhanced intensity (*blues*) indicates a short-lived particle that produced pions of a specific energy. The dark blue bands at 2.2 GeV^2 correspond to a particle dubbed $f_0(1500)$, which is possibly a glueball (1 GeV = 10^9 electron volts).

In early 1999 accelerators that produce reams of *B* mesons will start taking data in the U.S. and in Japan [see "The Asymmetry between Matter and Antimatter," by Helen R. Quinn and Michael S. Witherell; SCIENTIFIC AMERICAN, October]. Among their other uses, these mesons offer a promising new avenue for generating glueballs and hybrids.

Precision experiments dedicated to the search will also begin next year at Thomas Jefferson National Accelerator Facility in Newport News, Va. And in 2001 researchers at CERN will join in. One of these experiments will, we fondly hope, upturn unambiguous evidence of unadulterated glue. SA

The Authors

FRANK E. CLOSE and PHILIP R. PAGE began collaborating on gluon excitations in 1992 at the University of Oxford, where Close was Page's thesis supervisor. Close is currently head of theoretical particle physics at Rutherford Appleton Laboratory near Oxford and visiting professor at the University of Birmingham in England. He is also vice president of the British Association for the Advancement of Science. Page spent a year at Thomas Jefferson National Accelerator Facility in Newport News, Va., before moving to Los Alamos National Laboratory, where he is a staff scientist.

Further Reading

THE QUARK STRUCTURE OF MATTER. Frank Close in *The New Physics*. Edited by Paul Davies. Cambridge University Press, 1989.

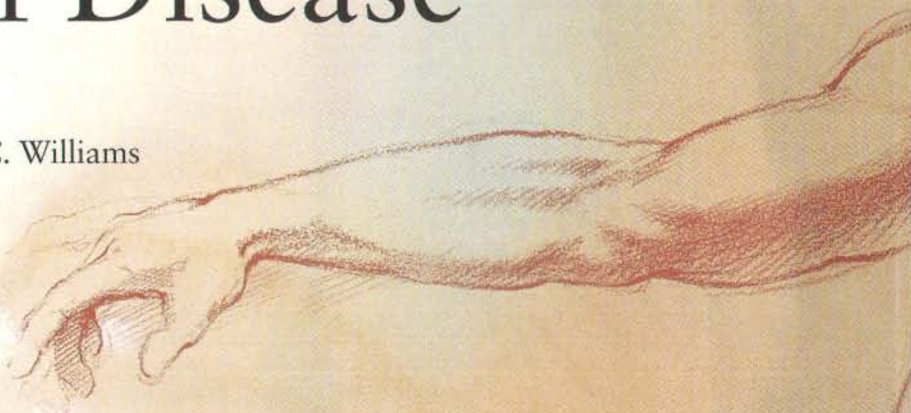
GLUEBALLS AND HYBRIDS: NEW STATES OF MATTER. F. E. Close in *Contemporary Physics*, Vol. 38, No. 1, pages 1–12; January 1997.

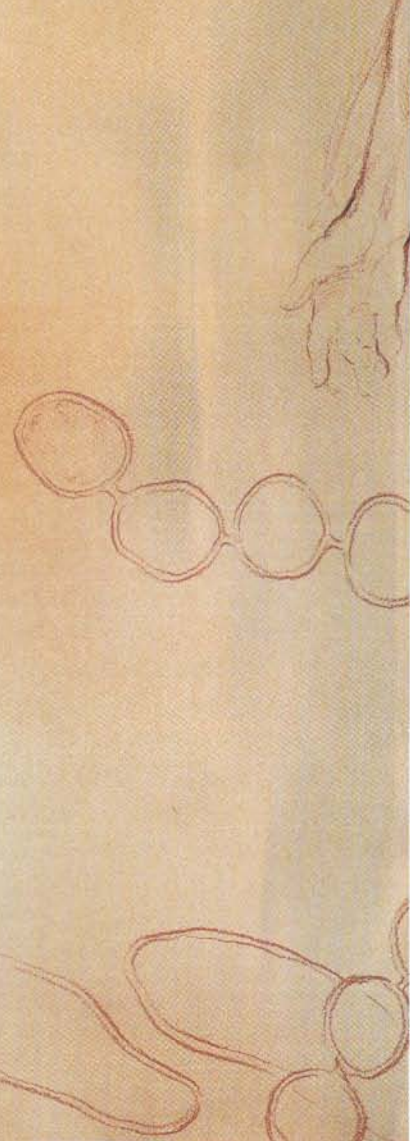
PARTICLE PHYSICS: EXOTIC MESON PLAYS HARD-TO-GET. F. E. Close in *Nature*, Vol. 389, pages 230–231; September 18, 1997.

The principles of evolution by natural selection are finally beginning to inform medicine

Evolution and the Origins of Disease

by Randolph M. Nesse and George C. Williams

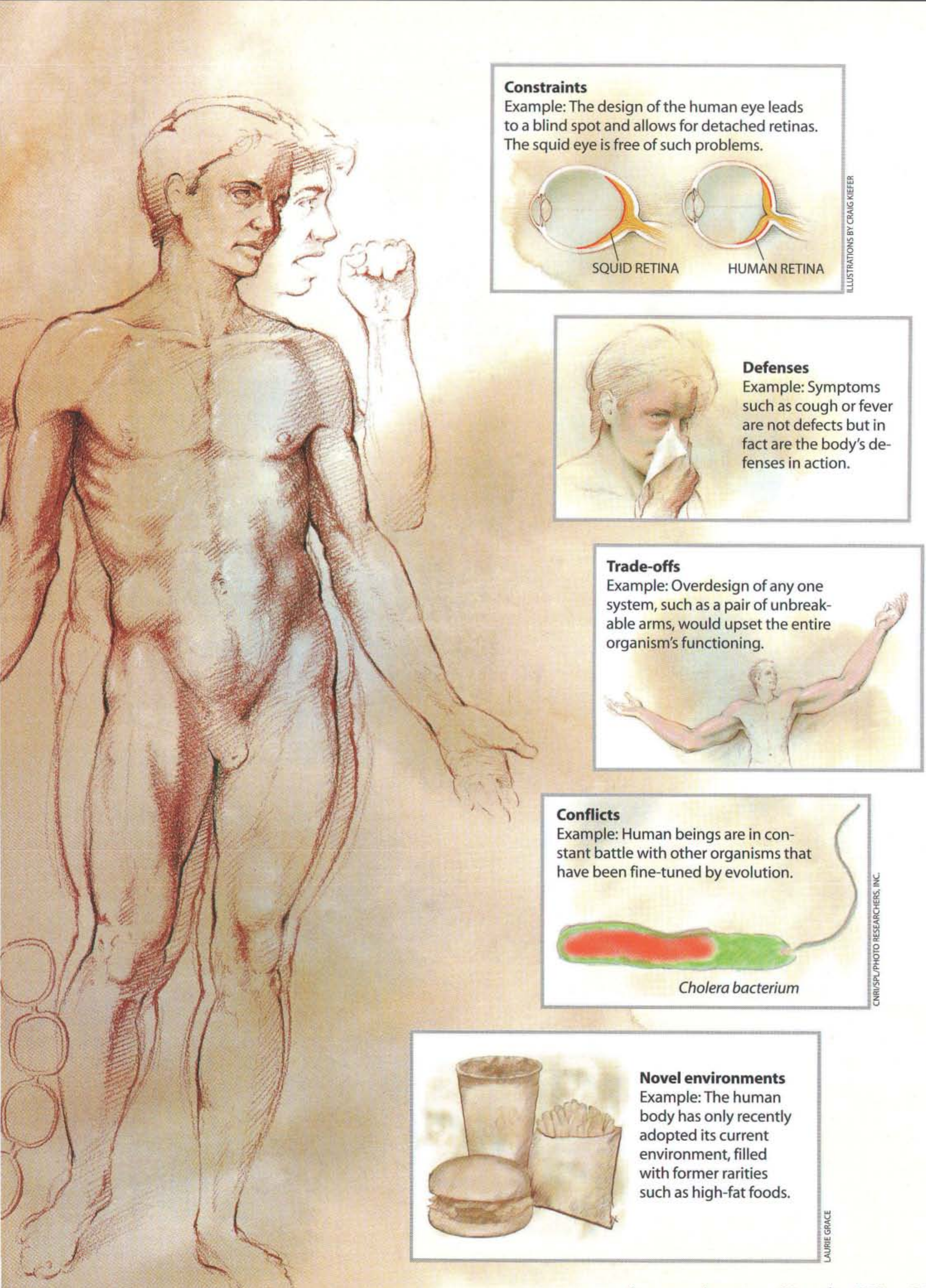




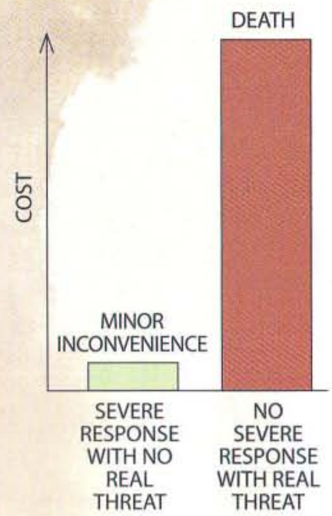
Thoughtful contemplation of the human body elicits awe—in equal measure with perplexity. The eye, for instance, has long been an object of wonder, with the clear, living tissue of the cornea curving just the right amount, the iris adjusting to brightness and the lens to distance, so that the optimal quantity of light focuses exactly on the surface of the retina. Admiration of such apparent perfection soon gives way, however, to consternation. Contrary to any sensible design, blood vessels and nerves traverse the inside of the retina, creating a blind spot at their point of exit.

The body is a bundle of such jarring contradictions. For each exquisite heart valve, we have a wisdom tooth. Strands of DNA direct the development of the 10 trillion cells that make up a human adult but then permit his or her steady deterioration and eventual death. Our immune system can identify and destroy a million kinds of foreign matter, yet many bacteria can still kill us. These contradictions make it appear as if the body was designed by a team of superb engineers with occasional interventions by Rube Goldberg.

In fact, such seeming incongruities make sense but only when we investigate the origins of the body's vulnerabilities while keeping in mind the wise words of distinguished geneticist Theodosius Dobzhansky: "Nothing in biology makes sense except in the light of evolution." Evolutionary biology is, of course, the scientific foundation for all biology, and biology is the foundation for all medicine. To a surprising degree, however, evolutionary bi-


Constraints
Example: The design of the human eye leads to a blind spot and allows for detached retinas. The squid eye is free of such problems.
SQUID RETINA
HUMAN RETINA
ILLUSTRATIONS BY CRAIG KIEFER
Defenses
Example: Symptoms such as cough or fever are not defects but in fact are the body's defenses in action.
Trade-offs
Example: Overdesign of any one system, such as a pair of unbreakable arms, would upset the entire organism's functioning.
Conflicts
Example: Human beings are in constant battle with other organisms that have been fine-tuned by evolution.
Cholera bacterium
CNRI/SPL/PHOTO RESEARCHERS, INC.
Novel environments
Example: The human body has only recently adopted its current environment, filled with former rarities such as high-fat foods.
LAURIE GRACE

CRAIG KIEFER



JOHNNY JOHNSON

MUCH SUFFERING is unnecessary but inevitable because of the smoke-detector nature of our defenses. The cost of a false alarm—a strong reaction such as vomiting in the absence of a true threat to life—is temporary unpleasantness. But the cost of no alarm in the presence of a true threat, such as a food toxin, could mean death. A lack of defensive response during pregnancy, for example, could kill the fetus.

ology is just now being recognized as a basic medical science. The enterprise of studying medical problems in an evolutionary context has been termed Darwinian medicine. Most medical research tries to explain the causes of an individual's disease and seeks therapies to cure or relieve deleterious conditions. These efforts are traditionally based on consideration of proximate issues, the straightforward study of the body's anatomic and physiological mechanisms as they currently exist. In contrast, Darwinian medicine asks why the body is designed in a way that makes us all vulnerable to problems like cancer, atherosclerosis, depression and choking, thus offering a broader context in which to conduct research.

The evolutionary explanations for the body's flaws fall into surprisingly few categories. First, some discomforting conditions, such as pain, fever, cough, vomiting and anxiety, are actually neither diseases nor design defects but rather are evolved defenses. Second, conflicts with other organisms—*Escherichia coli* or crocodiles, for instance—are a fact of life. Third, some circumstances, such as the ready availability of dietary fats, are so recent that natural selection has not yet had a chance to deal with them. Fourth, the body may fall victim to tradeoffs between a trait's benefits and its costs; a textbook example is the sickle cell gene, which also protects against malaria. Finally, the process of natural selection is constrained in ways that leave us with suboptimal design features, as in the case of the mammalian eye.

Evolved Defenses

Perhaps the most obviously useful defense mechanism is coughing; people who cannot clear foreign matter from their lungs are likely to die from pneumonia. The capacity for pain is also certainly beneficial. The rare individuals who cannot feel pain fail even to experience discomfort from staying in the same position for long periods. Their unnatural stillness impairs the blood supply to their joints, which then deteriorate. Such pain-free people usually die by early adulthood from tissue damage and infections. Cough or pain is usually interpreted as disease or trauma but is actually part of the solution rather than the problem. These defensive capabilities, shaped by natural selection, are kept in reserve until needed.

Less widely recognized as defenses are fever, nausea, vomiting, diarrhea, anxiety, fatigue, sneezing and inflammation. Even some physicians remain unaware of fever's utility. No mere increase in metabolic rate, fever is a carefully regulated rise in the set point of the body's thermostat. The higher body temperature facilitates the destruction of pathogens. Work by Matthew J. Kluger of the Lovelace Institute in Albuquerque, N.M., has shown that even cold-blooded lizards, when infected, move to warmer places until their bodies are several degrees above their usual temperature. If prevented from moving to the warm part of their cage, they are at increased risk of death from the infection. In a similar study by Evelyn Satinoff of the University of Delaware, elderly rats, who can no longer achieve the high fevers of their younger lab companions, also instinctively sought hotter environments when challenged by infection.

A reduced level of iron in the blood is another misunderstood defense mechanism. People suffering from chronic infection often have decreased levels of blood iron. Although such low iron is sometimes blamed for the illness, it actually is a protective response: during infection, iron is sequestered in the liver, which prevents invading bacteria from getting adequate supplies of this vital element.

Morning sickness has long been considered an unfortunate side effect of pregnancy. The nausea, however, coincides with the period of rapid tissue differentiation of the fetus, when development is most vulnerable to interference by toxins. And nauseated women tend to restrict their intake of strong-tasting, potentially harmful substances. These observations led independent researcher Margie Profet to hypothesize that the nausea of pregnancy is an adaptation whereby the mother protects the fetus from exposure to toxins. Profet tested this idea by examining pregnancy outcomes. Sure enough, women with more nausea were less likely to suffer miscarriages. (This evidence supports the hypothesis but is hardly conclusive. If Profet is correct, further research should discover that pregnant females of many species show changes in food preferences. Her theory also predicts an increase in birth defects among offspring of women who have little or no morning sickness and thus eat a wider variety of foods during pregnancy.)

Another common condition, anxiety, obviously originated as a defense in dangerous situations by promoting escape and avoidance. A 1992 study by Lee A. Dugatkin of the University of Louisville evaluated the benefits of fear in guppies. He grouped them as timid, ordinary or bold, depending on their reaction to the presence of smallmouth

bass. The timid hid, the ordinary simply swam away, and the bold maintained their ground and eyed the bass. Each guppy group was then left alone in a tank with a bass. After 60 hours, 40 percent of the timid guppies had survived, as had only 15 percent of the ordinary fish. The entire complement of bold guppies, on the other hand, wound up aiding the transmission of bass genes rather than their own.

Selection for genes promoting anxious behaviors implies that there should be people who experience too much anxiety, and indeed there are. There should also be hypophobic individuals who have insufficient anxiety, either because of genetic tendencies or antianxiety drugs. The exact nature and frequency of such a syndrome is an open question, as few people come to psychiatrists complaining of insufficient apprehension. But if sought, the pathologically nonanxious may be found in emergency rooms, jails and unemployment lines.

The utility of common and unpleasant conditions such as diarrhea, fever and anxiety is not intuitive. If natural selection shapes the mechanisms that regulate defensive responses, how can people get away with using drugs to block these defenses without doing their bodies obvious harm? Part of the answer is that we do, in fact, sometimes do ourselves a disservice by disrupting defenses.

Herbert L. DuPont of the University of Texas at Houston and Richard B. Hornick of Orlando Regional Medical Center studied the diarrhea caused by *Shigella* infection and found that people who took antidiarrhea drugs stayed sick longer and were more likely to have complications than those who took a placebo. In another example, Eugene D. Weinberg of Indiana University has documented that well-intentioned attempts to correct perceived iron deficiencies have led to increases in infectious disease, especially amebiasis, in parts of Africa. Although the iron in most oral supplements is unlikely to make much difference in otherwise healthy people with everyday infections, it can severely harm those who are infected and malnourished. Such people cannot make enough protein to bind the iron, leaving it free for use by infectious agents.

On the morning-sickness front, an antinausea drug was recently blamed for birth defects. It appears that no consideration was given to the possibility that the drug itself might be harmless to the fetus but could still be associated with birth defects, by interfering with the mother's defensive nausea.

Another obstacle to perceiving the benefits of defenses arises from the observation that many individuals regularly experience seemingly worthless reactions of anxiety, pain, fever, diarrhea or nausea. The explanation requires an analysis of the regulation of defensive responses in terms of signal-detection theory. A circulating toxin may come from something in the stomach. An organism can expel it by vomiting, but only at a price. The cost of a false alarm—vomiting when no toxin is truly present—is only a few calories. But the penalty for a single missed authentic alarm—failure to vomit when confronted with a toxin—may be death.

Natural selection therefore tends to shape regulation mechanisms with hair triggers, following what we call the smoke-detector principle. A smoke alarm that will reliably wake a sleeping family in the event of any fire will necessarily give a false alarm every time the toast burns. The price of the human body's numerous "smoke alarms" is much suffering that is completely normal but in most instances unnecessary. This principle also explains why blocking defenses is so often free of tragic consequences. Because most defensive reactions occur in response to insignificant threats, interference is usually harmless; the vast majority of alarms that are stopped by removing the battery from the smoke alarm are false ones, so this strategy may seem reasonable. Until, that is, a real fire occurs.

Conflicts with Other Organisms

Natural selection is unable to provide us with perfect protection against all pathogens, because they tend to evolve much faster than humans do. *E. coli,* for example, with its rapid rates of reproduction, has as much opportunity for mutation and selection in one day as humanity gets in a millennium. And our defenses, whether natural or artificial, make for potent selection forces. Pathogens either quickly evolve a counterdefense or become extinct. Amherst College biologist Paul W. Ewald has suggested classifying phenomena associated with infection according to whether they benefit the host, the pathogen, both or neither. Consider the runny nose associated with a cold. Nasal mucous secretion could expel intruders,

Evolution of Virulence

Changes in virulence relate to the life history of the infectious agent and its mode of transmission. As elucidated by Paul W. Ewald of Amherst College, infection requiring direct contact will ordinarily drive a pathogen toward a state of lowered virulence, because the host must remain mobile enough to interact with others. But intermediaries that spread disease-causing agents, even from totally incapacitated hosts, can cause a change toward more virulence. Behavioral choices, such as safer sex, can also alter the makeup of the pathogen.

Intermediary disease vectors
(Mosquitoes, health care workers' hands, unsanitary water supplies)
Unprotected and/or promiscuous sex

Casual human-to-human transmission
(Sneezing, coughing, touch)
Protected and/or monogamous sex

CRAIG KIEFER

speed the pathogen's transmission to new hosts or both [see "The Evolution of Virulence," by Paul W. Ewald; Scientific American, April 1993]. Answers could come from studies examining whether blocking nasal secretions shortens or prolongs illness, but few such studies have been done.

Humanity won huge battles in the war against pathogens with the development of antibiotics and vaccines. Our victories were so rapid and seemingly complete that in 1969 U.S. Surgeon General William H. Stewart said that it was "time to close the book on infectious disease." But the enemy, and the power of natural selection, had been underestimated. The sober reality is that pathogens apparently can adapt to every chemical researchers develop. ("The war has been won," one scientist more recently quipped. "By the other side.")

Antibiotic resistance is a classic demonstration of natural selection. Bacteria that happen to have genes that allow them to prosper despite the presence of an antibiotic reproduce faster than others, and so the genes that confer resistance spread quickly. As shown by Nobel laureate Joshua Lederberg of the Rockefeller University, they can even jump to different species of bacteria, borne on bits of infectious DNA. Today some strains of tuberculosis in New York City are resistant to all three main antibiotic treatments; patients with those strains have no better chance of surviving than did TB patients a century ago. Stephen S. Morse of Columbia University notes that the multidrug-resistant strain that has spread throughout the East Coast may have originated in a homeless shelter across the street from Columbia-Presbyterian Medical Center. Such a phenomenon would indeed be predicted in an environment where fierce selection pressure quickly weeds out less hardy strains. The surviving bacilli have been bred for resistance.

Many people, including some physicians and scientists, still believe the outdated theory that pathogens necessarily become benign after long association with hosts. Superficially, this makes sense. An organism that kills rapidly may never get to a new host, so natural selection would seem to favor lower virulence. Syphilis, for instance, was a highly virulent disease when it first arrived in Europe, but as the centuries passed it became steadily more mild. The virulence of a pathogen is, however, a life history trait that can increase as well as decrease, depending on which option is more advantageous to its genes.

For agents of disease that are spread directly from person to person, low virulence tends to be beneficial, as it allows the host to remain active and in contact with other potential hosts. But some diseases, like malaria, are transmitted just as well—or better—by the incapacitated. For such pathogens, which usually rely on intermediate vectors like mosquitoes, high virulence can give a selective advantage. This principle has direct implications for infection control in hospitals, where health care workers' hands can be vectors that lead to selection for more virulent strains.

In the case of cholera, public water supplies play the mosquitoes' role. When water for drinking and bathing is contaminated by waste from immobilized patients, selection tends to increase virulence, because more diarrhea enhances the spread of the organism even if individual hosts quickly die. But, as Ewald has shown, when sanitation improves, selection acts against classical *Vibrio cholerae* bacteria in favor of the more benign El Tor biotype. Under these conditions, a dead host is a dead end. But a less ill and more mobile host, able to infect many others over a much longer time, is an effective vehicle for a pathogen of lower virulence. In another example, better sanitation leads to displacement of the aggressive *Shigella flexneri* by the more benign *S. sonnei*.

Such considerations may be relevant for public policy. Evolutionary theory predicts that clean needles and the encouragement of safe sex will do more than save numerous individuals from HIV infection. If humanity's behavior itself slows HIV transmission rates, strains that do not soon kill their hosts have the long-term survival advantage over the more virulent viruses that then die with their hosts, denied the opportunity to spread. Our collective choices can change the very nature of HIV.

Conflicts with other organisms are not limited to pathogens. In times past, humans were at great risk from predators looking for a meal. Except in a few places, large carnivores now pose no threat to humans. People are in more danger today from smaller organisms' defenses, such as the venoms of spiders and snakes. Ironically, our fears of small creatures, in the form of phobias, probably cause more harm than any interactions with those organisms do. Far more dangerous than predators or poisoners are other members of our own species. We attack each other not to get meat but to get mates, territory and other re-

New Environments Bring New Health Threats

Common Threats to Health from 20,000 B.C. to Modern Times

Accidents

Starvation

Predation

Infectious diseases

Common Threats to Health Today (in Technologically Advanced Cultures)

Heart attack, stroke and other complications of atherosclerosis

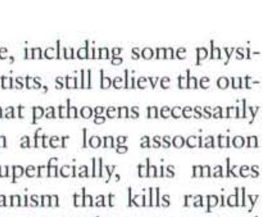
Cancer

Other chronic diseases associated with lifestyle and longevity

Noninsulin-dependent diabetes

Obesity

New infectious diseases



sources. Violent conflicts between individuals are overwhelmingly between young men in competition and give rise to organizations to advance these aims. Armies, again usually composed of young men, serve similar objectives, at huge cost.

Even the most intimate human relationships give rise to conflicts having medical implications. The reproductive interests of a mother and her infant, for instance, may seem congruent at first but soon diverge. As noted by biologist Robert L. Trivers in a now classic 1974 paper, when her child is a few years old, the mother's genetic interests may be best served by becoming pregnant again, whereas her offspring benefits from continuing to nurse. Even in the womb there is contention. From the mother's vantage point, the optimal size of a fetus is a bit smaller than that which would best serve the fetus and the father. This discord, according to David Haig of Harvard University, gives rise to an arms race between fetus and mother over her levels of blood pressure and blood sugar, sometimes resulting in hypertension and diabetes during pregnancy.

Coping with Novelty

Making rounds in any modern hospital provides sad testimony to the prevalence of diseases humanity has brought on itself. Heart attacks, for example, result mainly from atherosclerosis, a problem that became widespread only in this century and that remains rare among hunter-gatherers. Epidemiological research furnishes the information that should help us prevent heart attacks: limit fat intake, eat lots of vegetables, and exercise hard each day. But hamburger chains proliferate, diet foods languish on the shelves, and exercise machines serve as expensive clothing hangers throughout the land. The proportion of overweight Americans is one third and rising. We all know what is good for us. Why do so many of us continue to make unhealthy choices?

Our poor decisions about diet and exercise are made by brains shaped to cope with an environment substantially different from the one our species now inhabits. On the African savanna, where the modern human design was fine-tuned, fat, salt and sugar were scarce and precious. Individuals who had a tendency to consume large amounts of fat when given the rare opportunity had a selective advantage. They were more likely to survive famines that killed their thinner companions. And we, their descendants, still carry those urges for foodstuffs that today are anything but scarce. These evolved desires—inflamed by advertisements from competing food corporations that themselves survive by selling us more of whatever we want to buy—easily defeat our intellect and willpower. How ironic that humanity worked for centuries to create environments that are almost literally flowing with milk and honey, only to see our success responsible for much modern disease and untimely death.

Increasingly, people also have easy access to many kinds of drugs, especially alcohol and tobacco, that are responsible for a huge proportion of disease, health care costs and premature death. Although individuals have always used psychoactive substances, widespread problems materialized only following another environmental novelty: the ready availability of concentrated drugs and new, direct routes of administration, especially injection. Most of these substances, including nicotine, cocaine and opium, are products of natural selection that evolved to protect plants from insects. Because humans share a common evolutionary heritage with insects, many of these substances also affect our nervous system.

This perspective suggests that it is not just defective individuals or disordered societies that are vulnerable to the dangers of psychoactive drugs; all of us are susceptible because drugs and our biochemistry have a long history of interaction. Understanding the details of that interaction, which is the focus of much current research from both a proximate and evolutionary perspective, may well lead to better treatments for addiction.

The relatively recent and rapid increase in breast cancer must be the result in large part of changing environments and ways of life, with only a few cases resulting solely from genetic abnormalities. Boyd Eaton and his colleagues at Emory University reported that the rate of breast cancer in today's "nonmodern" societies is only a tiny fraction of that in the U.S. They hypothesize that the amount of time between menarche and first pregnancy is a crucial risk factor, as is the related issue of total lifetime number of menstrual cycles. In hunter-gatherers, menarche occurs at about age 15 or later, followed within a few years by pregnancy and two or three years of nursing, then by another pregnancy soon after. Only between the end of nursing and the next pregnancy will the woman menstruate and thus experience the high levels of hormones that may adversely affect breast cells.

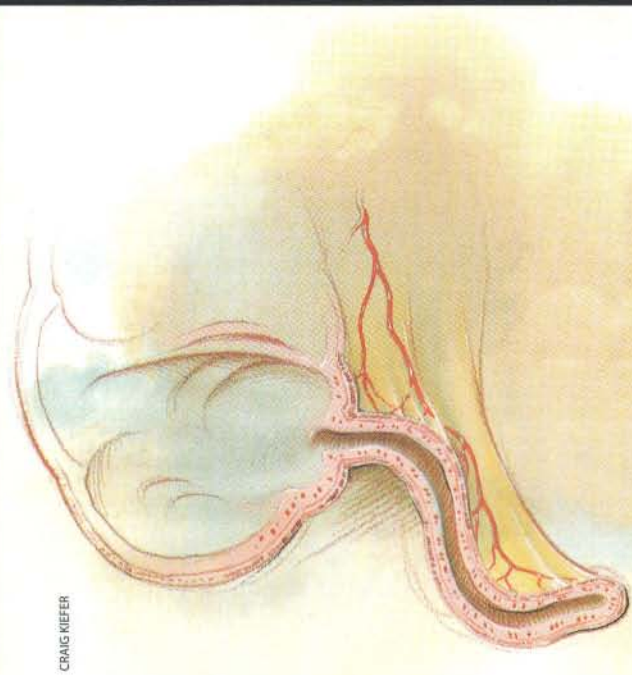
In modern societies, in contrast, menarche occurs at age 12 or 13—probably at least in part because of a fat intake sufficient to allow an extremely young woman to nourish a fetus—and the first pregnancy may be decades later or never. A female hunter-gatherer may have a total of 150 menstrual cycles, whereas the average woman in modern societies has 400 or more. Although few would suggest that women should become pregnant in their teens to prevent breast cancer later, early administration of a burst of hormones to simulate pregnancy may reduce the risk. Trials to test this idea are now under way at the University of California at San Diego.

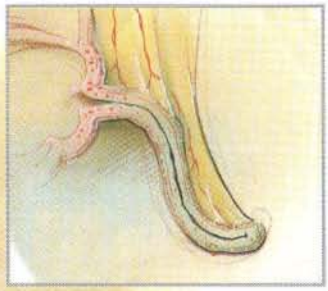
Trade-offs and Constraints

Compromise is inherent in every adaptation. Arm bones three times their current thickness would almost never break, but *Homo sapiens* would be lumbering creatures on a never-ending quest for calcium. More sensitive ears might sometimes be useful, but we would be distracted by the noise of air molecules banging into our eardrums.

Such trade-offs also exist at the genetic level. If a mutation offers a net reproductive advantage, it will tend to increase in frequency in a population even if it causes vulnerability to disease. People with two copies of the sickle cell gene, for example, suffer terrible pain and die young. People with two copies of the "normal" gene are at high risk of death from malaria. But individuals with one of each are protected from both malaria and sickle cell disease. Where malaria is prevalent, such people are fitter, in the Darwinian sense, than members of either other group. So even though the sickle cell gene causes disease, it is selected for where malaria persists. Which is the "healthy" allele in this environment? The question has no answer. There is no one normal human genome—there are only genes.

Many other genes that cause disease must also have offered benefits, at least in some environments, or they would not be so common. Because cystic fibrosis (CF) kills one out of 2,500 Caucasians, the responsible genes would ap-





APPENDIX is most likely here to stay. Evolutionary pressure selects against the smaller appendix (*above*), because inflammation and swelling can cut off its cleansing blood supply, making infections more life-threatening. Larger appendices are thus actually selected for.

CRAIG KIEFER

pear to be at great risk of being eliminated from the gene pool. And yet they endure. For years, researchers mused that the CF gene, like the sickle cell gene, probably conferred some advantage. Recently a study by Gerald B. Pier of Harvard Medical School and his colleagues gave substance to this informed speculation: having one copy of the CF gene appears to decrease the chances of the bearer acquiring a typhoid fever infection, which once had a 15 percent mortality.

Aging may be the ultimate example of a genetic trade-off. In 1957 one of us (Williams) suggested that genes that cause aging and eventual death could nonetheless be selected for if they had other effects that gave an advantage in youth, when the force of selection is stronger. For instance, a hypothetical gene that governs calcium metabolism so that bones heal quickly but that also happens to cause the steady deposition of calcium in arterial walls might well be selected for even though it kills some older people. The influence of such pleiotropic genes (those having multiple effects) has been seen in fruit flies and flour beetles, but no specific example has yet been found in humans. Gout, however, is of particular interest, because it arises when a potent antioxidant, uric acid, forms crystals that precipitate out of fluid in joints. Antioxidants have antiaging effects, and plasma levels of uric acid in different species of primates are closely correlated with average adult life span. Perhaps high levels of uric acid benefit most humans by slowing tissue aging, while a few pay the price with gout.

Other examples are more likely to contribute to more rapid aging. For instance, strong immune defenses protect us from infection but also inflict continuous, low-level tissue damage. It is also possible, of course, that most genes that cause aging have no benefit at any age—they simply never decreased reproductive fitness enough in the natural environment to be selected against. Nevertheless, over the next decade research will surely identify specific genes that accelerate senescence, and researchers will soon thereafter gain the means to interfere with their actions or even change them. Before we tinker, however, we should determine whether these actions have benefits early in life.

Because evolution can take place only in the direction of time's arrow, an organism's design is constrained by structures already in place. As noted, the vertebrate eye is arranged backward. The squid eye, in contrast, is free from this defect, with vessels and nerves running on the outside, penetrating where necessary and pinning down the retina so it cannot detach. The human eye's flaw results from simple bad luck; hundreds of millions of years ago, the layer of cells that happened to become sensitive to light in our ancestors was positioned differently from the corresponding layer in ancestors of squids. The two designs evolved along separate tracks, and there is no going back.

Such path dependence also explains why the simple act of swallowing can be life-threatening. Our respiratory and food passages intersect because in an early lungfish ancestor the air opening for breathing at the surface was understandably located at the top of the snout and led into a common space shared by the food passageway. Because natural selection cannot start from scratch, humans are stuck with the possibility that food will clog the opening to our lungs.

The path of natural selection can even lead to a potentially fatal cul-de-sac, as in the case of the appendix, that vestige of a cavity that our ancestors employed in digestion. Because it no longer performs that function, and as it can kill when infected, the expectation might be that natural selection would have eliminated it. The reality is more complex. Appendicitis results when inflammation causes swelling, which compresses the artery supplying blood to the appendix. Blood flow protects against bacterial growth, so any reduction aids infection, which creates more swelling. If the blood supply is cut off completely, bacteria have free rein until the appendix bursts. A slender appendix is especially susceptible to this chain of events, so appendicitis may, paradoxically, apply the selective pressure that maintains a large appendix. Far from arguing that everything in the body is perfect, an evolutionary analysis reveals that we live with some very unfortunate legacies and that some vulnerabilities may even be actively maintained by the force of natural selection.

Evolution of Darwinian Medicine

Despite the power of the Darwinian paradigm, evolutionary biology is just now being recognized as a basic science essential for medicine. Most diseases decrease fitness, so it would seem that natural selection could explain only health, not disease. A Darwinian approach makes sense only when the object of explanation is changed from diseases to the traits that make us vulnerable to diseases. The assumption that natural selection maximizes health also is incorrect—selection maximizes the reproductive success of genes. Those genes that make bodies having superior reproductive success will become more common, even if they compromise the individual's health in the end.

Finally, history and misunderstanding have presented obstacles to the acceptance of Darwinian medicine. An evolutionary approach to functional analysis can appear akin to naive teleology or vitalism, errors banished only recently, and with great effort, from medical thinking. And, of course, whenever evolution and medicine are mentioned together, the specter of eugenics arises.

Selected Principles of Darwinian Medicine

A Darwinian approach to medical practice leads to a shift in perspective. The following principles provide a foundation for considering health and disease in an evolutionary context:

DEFENSES and **DEFECTS** are two fundamentally different manifestations of disease

BLOCKING defenses has costs as well as benefits

Because natural selection shapes defense regulation according to the **SMOKE-DETECTOR PRINCIPLE**, much defensive expression and associated suffering are unnecessary in the individual instance

Modern epidemics are most likely to arise from the mismatch between **PHYSIOLOGICAL DESIGN** of our bodies and **NOVEL ASPECTS** of our environment

Our **DESIRES**, shaped in the ancestral environment to lead us to actions that tended to maximize reproductive success, now often lead us to disease and early death

The body is a bundle of **COMPROMISES**

There is no such thing as "the **NORMAL** body"

There is no such thing as "the **NORMAL** human genome"

Some **GENES** that cause disease may also have benefits, and others are quirks that cause disease only when they interact with novel environmental factors

GENETIC SELF-INTEREST will drive an individual's actions, even at the expense of the health and longevity of the individual created by those genes

VIRULENCE is a trait of the pathogen that can increase as well as decrease

SYMPTOMS of infection can benefit the pathogen, the host, both or neither

Disease is **INEVITABLE** because of the way that organisms are shaped by evolution

Each disease needs a **PROXIMATE EXPLANATION** of why some people get it and others don't, as well as an **EVOLUTIONARY EXPLANATION** of why members of the species are vulnerable to it

Diseases are not products of natural selection, but most of the **VULNERABILITIES** that lead to disease are shaped by the process of natural selection

Aging is better viewed as a **TRADE-OFF** than a disease

Specific clinical recommendations must be based on **CLINICAL STUDIES**; clinical interventions based only on theory are not scientifically grounded and may cause harm

Discoveries made through a Darwinian view of how all human bodies are alike in their vulnerability to disease will offer great benefits for individuals, but such insights do not imply that we can or should make any attempt to improve the species. If anything, this approach cautions that apparent genetic defects may have unrecognized adaptive significance, that a single "normal" genome is nonexistent and that notions of "normality" tend to be simplistic.

The systematic application of evolutionary biology to medicine is a new enterprise. Like biochemistry at the beginning of this century, Darwinian medicine very likely will need to develop in several incubators before it can prove its power and utility. If it must progress only from the work of scholars without funding to gather data to test their ideas, it will take decades for the field to mature. Departments of evolutionary biology in medical schools would accelerate the process, but for the most part they do not yet exist. If funding agencies had review panels with evolutionary expertise, research would develop faster, but such panels remain to be created. We expect that they will.

The evolutionary viewpoint provides a deep connection between the states of disease and normal functioning and can integrate disparate avenues of medical research as well as suggest fresh and important areas of inquiry. Its utility and power will ultimately lead to recognition of evolutionary biology as a basic medical science.

The Authors

RANDOLPH M. NESSE and GEORGE C. WILLIAMS are the authors of the 1994 book *Why We Get Sick: The New Science of Darwinian Medicine.* Nesse received his medical degree from the University of Michigan Medical School in 1974. He is now professor of psychiatry at that institution and is director of the Evolution and Human Adaptation Program at the university's Institute for Social Research. Williams received his doctorate in 1955 from the University of California, Los Angeles, and quickly became one of the world's foremost evolutionary theorists. A member of the National Academy of Sciences, he is professor emeritus of ecology and evolution at the State University of New York at Stony Brook and edits the *Quarterly Review of Biology.*

Further Reading

Evolution of Infectious Disease. P. W. Ewald. Oxford University Press, 1994.

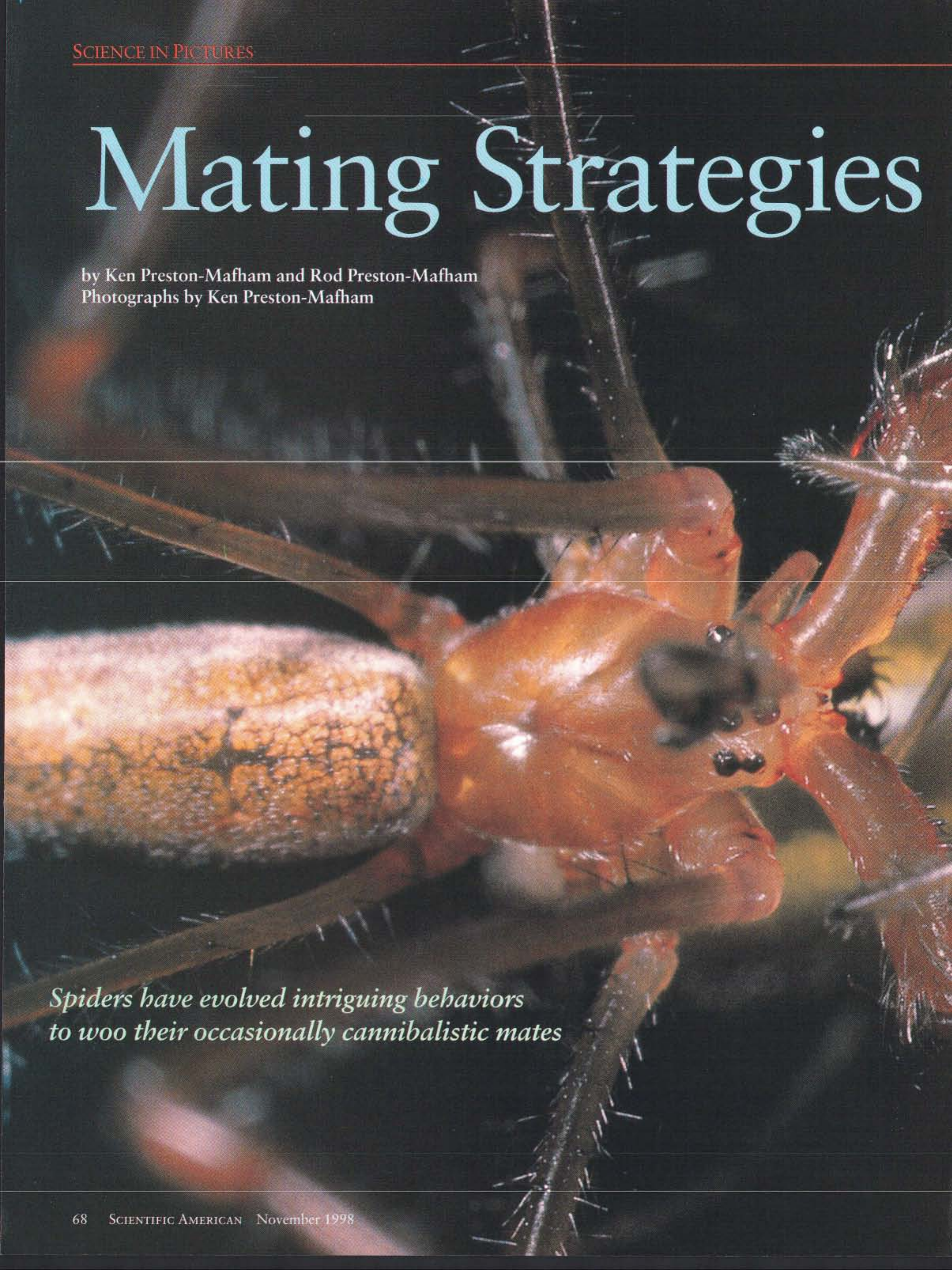
Darwinian Psychiatry. M. T. McGuire and A. Troisi. Harvard University Press, 1998.

Evolution in Health and Disease. Edited by S. Stearns. Oxford University Press, 1998.

Evolutionary Medicine. W. R. Trevathan et al. Oxford University Press (in press).

Mating Strategies

by Ken Preston-Mafham and Rod Preston-Mafham
Photographs by Ken Preston-Mafham



Spiders have evolved intriguing behaviors to woo their occasionally cannibalistic mates

# of Spiders

GRIPPING THE FEMALE'S FANGS to forestall a possible bite, a long-jawed orb-weaver (*Tetragnatha extensa*) male moves in to mate. The female (*at right*) is willing; else she would have warned off the male by bouncing on her web. Moreover, she greets the male with bared fangs that apparently signal not a threat but an enticement. Even so, mating will require about 15 minutes, and the male takes no chances in case she should get peckish. In spite of her total lack of aggression, he dives straight out of the web as soon as they have finished mating.

Range: Eurasia and northern North America; shown approximately 27 times life size

A spider's reproductive system bears little resemblance to that of any other group in the animal kingdom. From its head projects a pair of appendages called the pedipalps, sensory structures used primarily for tasting prospective prey. In male spiders the terminal segments of the "palps" are modified for introducing semen into a female.

The palpal organ, the basic copulatory structure of the male palp, can be likened to a simple bulb pipette. From a chamber that acts as a reservoir runs a narrow tube called the embolus with a pointed, open end through which semen passes. The male introduces the embolus into the female reproductive opening during mating. In higher families of spiders the palpal organ is surrounded by a complex set of sclerotized plates, hooks and spines. The females have simultaneously evolved a sclerotized structure, the epigyne, near the reproductive opening. The projections on the male palp will fit only into the epigyne of a female of the same species.

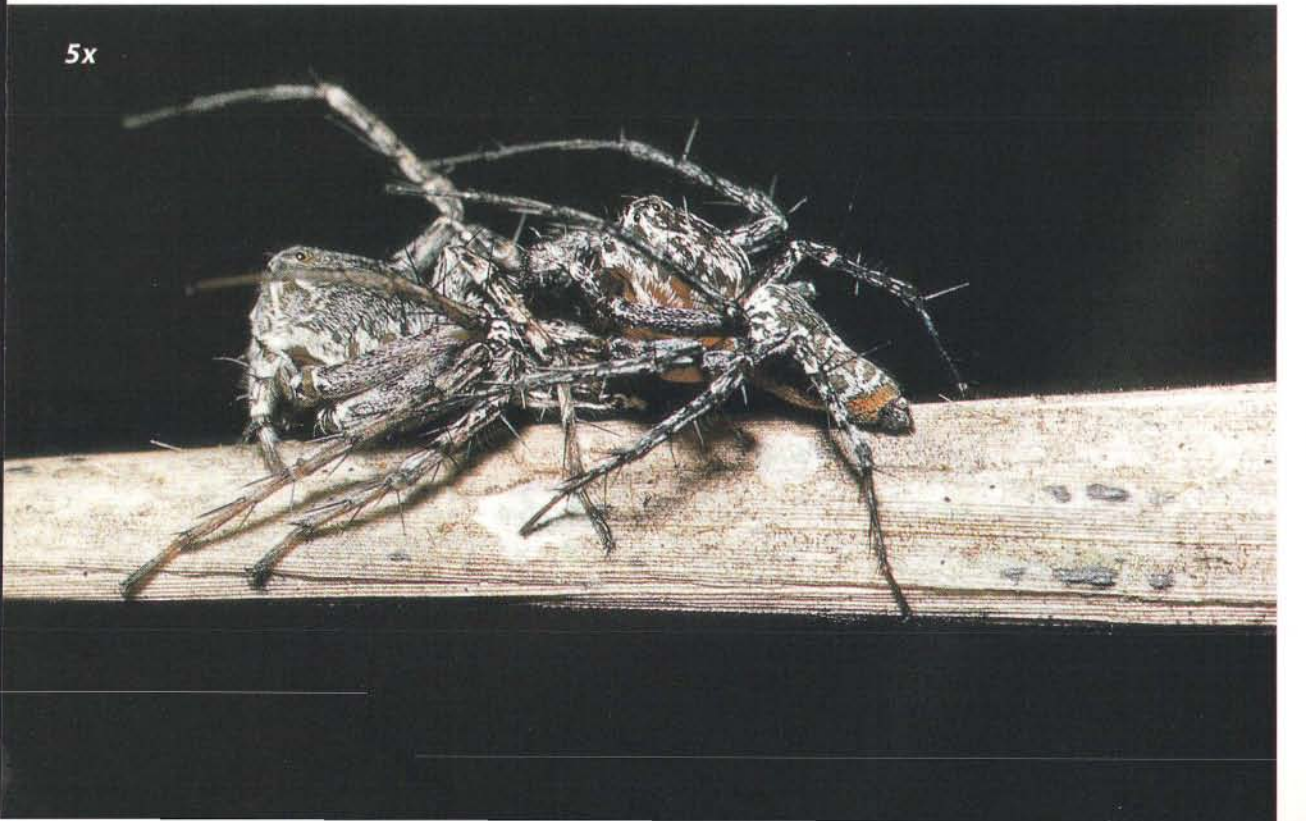
The palpal organs are not directly connected to the testes, so before mating can take place the male must somehow fill the palpal reservoirs with semen. First the male deposits a drop of semen onto silk from his spinnerets. This silk may be a single line stretched between a pair of legs, a simple, loose web held in a similar fashion or, in the higher spiders, a specially constructed web built onto an adjacent substrate. Then, holding the tip of the embolus against the semen, the male pumps it into the reservoir. In some species the male fills both palpal organs at the same time; in others, he fills only one, returning after a bout of copulation to prime the other. Once charged with semen, the male goes in search of a suitably receptive female. (In the family Linyphiidae, however, the males start sperm induction only when stimulated by contact with pheromones on the female web.)

Male web-building spiders would seem to face an insurmountable obstacle in finding mates—they have to search in three dimensions. They are, however, assisted by the females, which advertise the presence of their webs by releasing guidance pheromones. Female spiders that do not build permanent webs but wander around in search of prey always trail a single strand of silk dragline impregnated with pheromones. A wandering male of the same species coming across the dragline is thus able to follow it until he finds her. Females of the American fishing spider *Dolomedes triton* release their pheromones into the water near where they hunt, allowing male spiders to find them.

Spiders may also take part in some form of courtship. It can vary from the perfunctory, where the male just walks up

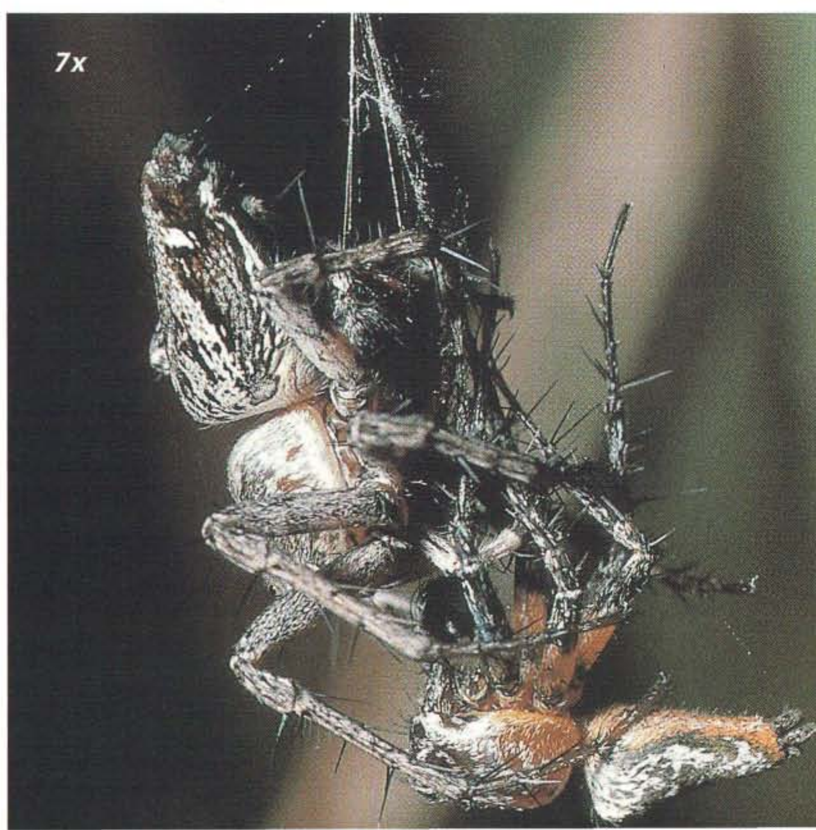
SPINNING A BRIDAL VEIL of silk, a male lynx spider (*Oxyopes schenkeli*) courts a female. First he strokes her back briefly with his front legs (*left*). In response, she walks away and drops off the edge of a leaf, dangling from her dragline. The male climbs down and begins to twirl her round and round, trailing silk across her body like a bobbin (*center*). When she is well wrapped, he tries to mate. Unfortunately, near the end of the mating season, when these pictures were taken, the female's cooperation can evaporate when the spinning stops. This one easily broke loose from her bonds and pinned the male in her fangs (*right*), then proceeded to feed on him for the next hour.

Range: Central Africa, mainly Congo and Uganda



COPULATING FROM DAWN TO DUSK may be a survival strategy for *Tylorida ventralis,* another long-jawed orb weaver. The female (*at left*) is extraordinarily cooperative: not only does she mate all day, every day, for a week or so with the same male but she often invites the first move. The behavior may be occasioned by the presence of several spider-hunting spiders in the region; having a partner halves her chances of becoming a meal.

Range: South Asia and eastward to New Guinea



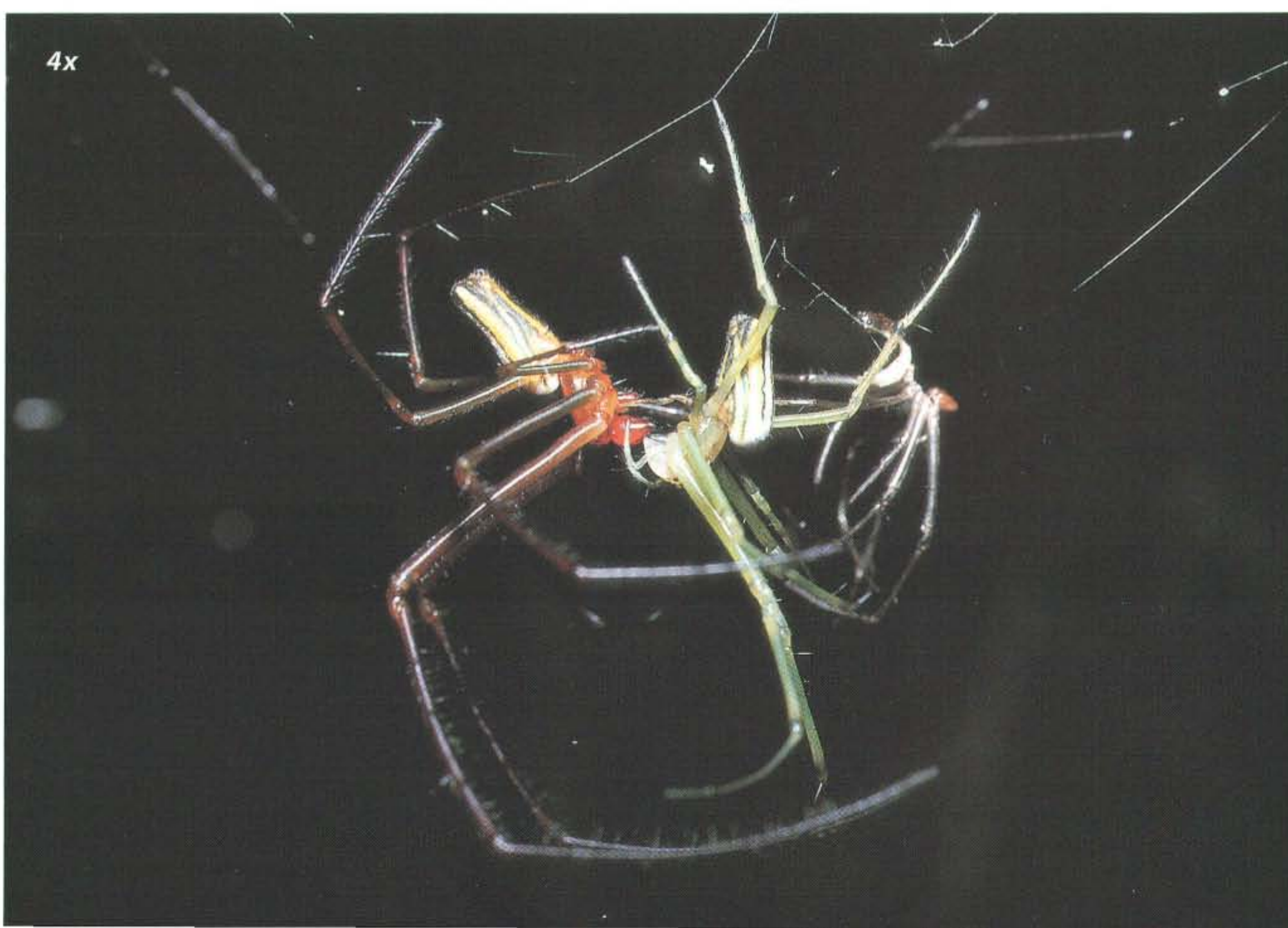
to the female and induces her to accept him, to the prolonged, where the male may approach and be rebuffed many times before ultimately getting to mate. In large-eyed, active hunters, such as wolf and jumping spiders, the males signal their intentions to the females with their sometimes colorfully marked palps. Some males physically subdue females before mating.

Mating can certainly be dangerous for some males. In a limited number of species—the widows in particular—the female often eats the male after copulation. In many spiders, however, males and females have jaws of the same size. Because the male is considerably more agile, the female may be at as much risk of being attacked by him as he is by her. Moreover, some female spiders are many times the size of the males, yet the latter seem quite happy to reside in the same web.

All in all, these odd mating routines have proved extremely successful for the spiders. SA

GUARDING a molting female (*right*), a male *Leucauge nigrovittata* ensures that he will mate with a virgin. When the newly hatched female has recovered her strength, she vibrates her body to invite him. After mating, he leaves, but she invites him back two more times. Their copulation was, however, interrupted by a second female, who rushed into the web, demanding to mate. The male tried to drive her off, but she refused to give up, and he ultimately consented to mate with her briefly before returning to the first female (*below*).

Range: Java, Sumatra and Borneo



OFFERING THE CORPSE of a rival, a common orb-weaver (*Meta segmentata*) male conducts the courtship ritual. Several males often fight for position in a female's web, leading to fatalities. This male (*at left*) has turned his victory to good account, using the loser's silk-wrapped body as a nuptial gift to the female. Normally he would wait for a fly to fall into the web.

Range: Europe, temperate Asia and Canada

LAYING DOWN LACE of fine silk, a male giant wood spider (*Nephila maculata*) treads on the back of the huge female. Although she generally ignores him, the female sometimes tries to flick him off. When she does succeed, she begins to undo his painstaking handiwork. Perhaps to avoid such interruptions, males tend to lay silk while the female is busy feeding. The strands, which are applied near the female's scent receptors, are saturated with pheromones and may convey a sexual message.

Range: South Asia to Japan and northern Australia

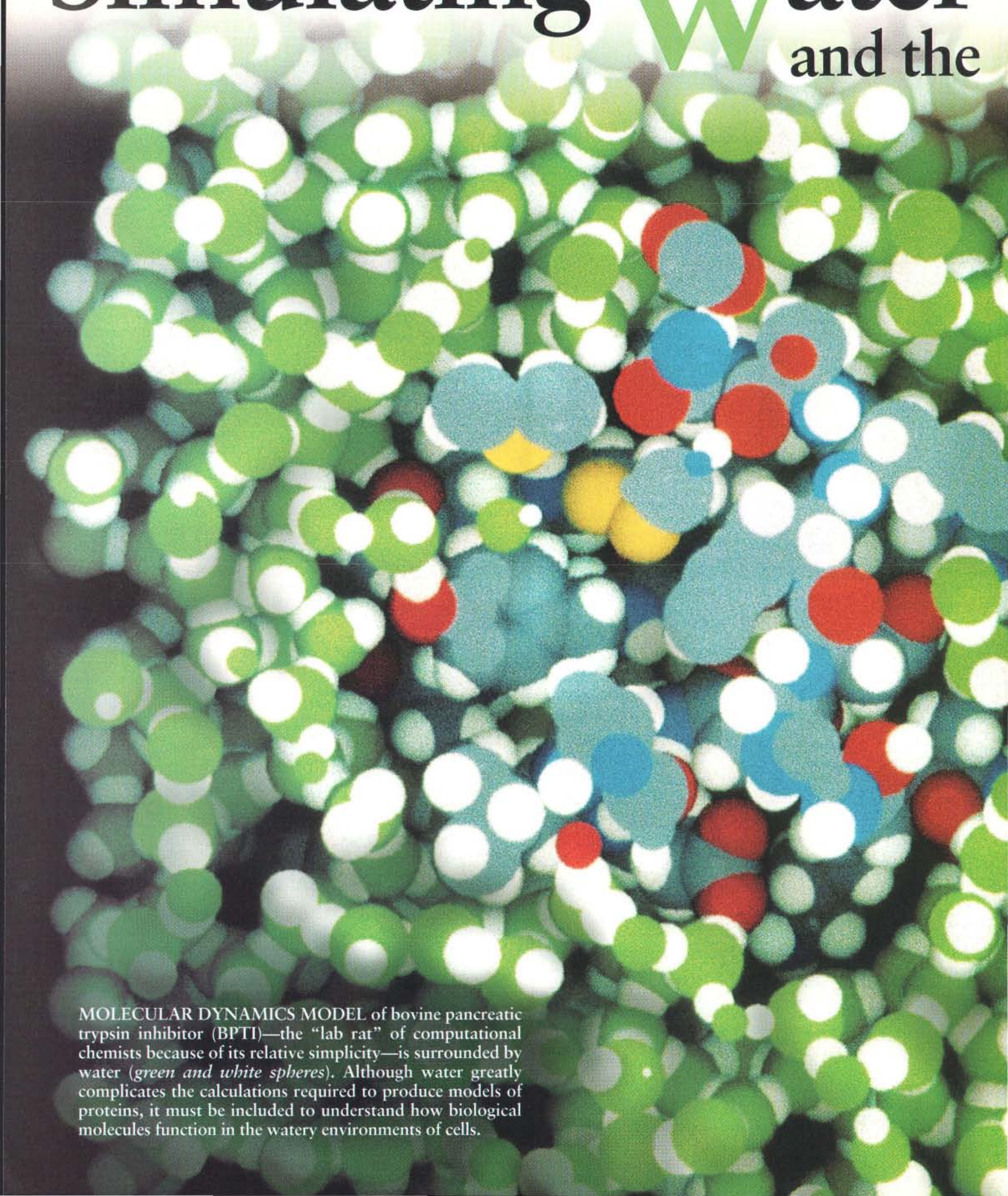
BINDING HIS MATE TO A LEAF, a male European crab spider (*Xysticus cristatus*) will walk all over the much larger but cooperative female. He then climbs beneath her and mates for about an hour, after which he strolls off. The female is able to disentangle herself with relative ease, indicating that the tether is not a safety device for the male. Its purpose is unknown.

Range: Europe, North Africa and much of Asia

The Authors

KEN PRESTON-MAFHAM and ROD PRESTON-MAFHAM are brothers and naturalists. Ken is a photographer who has captured in his lens more than 3,000 species in 40 countries, especially insects and spiders. He is the author or co-author of nearly 20 books. In recent years he has taken to original research, publishing papers on insect and spider behavior. Rod is trained as a zoologist and obtained his Ph.D. at the University of London. He has written a number of books and prepared programs for BBC radio, all on natural history. With his brother and his wife, Jean, he runs a wildlife picture library, Premaphotos Wildlife.

Simulating Water and the

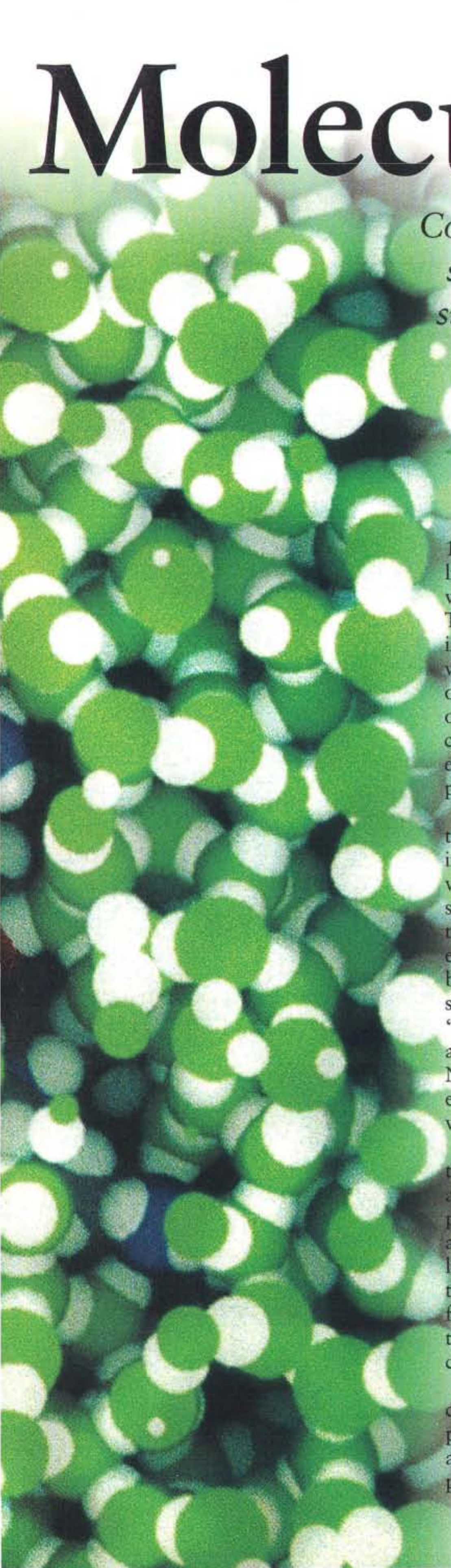


MOLECULAR DYNAMICS MODEL of bovine pancreatic trypsin inhibitor (BPTI)—the "lab rat" of computational chemists because of its relative simplicity—is surrounded by water (*green and white spheres*). Although water greatly complicates the calculations required to produce models of proteins, it must be included to understand how biological molecules function in the watery environments of cells.

Molecules of Life

Computer modeling reveals how water affects the structures and dynamics of biological molecules such as proteins, yielding clues to their functions

by Mark Gerstein and Michael Levitt



Water is cheap, if not free, in most places in the world. But during the summer of 1986, one of us (Levitt) spent half a million dollars on an amount of water that would scarcely wet the point of a pin. The money was not to buy the vanishingly small amount of water. Rather it was to pay for the roughly two weeks of processing time on a gigantic state-of-the-art supercomputer required to create a model of how the water affected the structure and movement of a particular protein.

The protein was bovine pancreatic trypsin inhibitor (BPTI), which is found in the pancreases of cattle. BPTI is a favorite subject of computer modelers simply because it is relatively small and therefore easier to study than most other proteins. It had been modeled before, by Martin Karplus of Harvard University and his colleagues in 1977, but only "in vacuo" (as if in a vacuum)—without any other molecules interacting with it. No one had visualized BPTI as it really exists in a living cell, with thousands of water molecules surrounding it.

The half a million dollars turned out to be well spent. Not only did Levitt and his colleague Ruth Sharon find the previous in vacuo model of BPTI to be a poor predictor of how the protein looked and behaved in the real world, the discovery helped to pave the way for computational chemists to simulate the structures of other biological molecules in their native, watery environs.

Today, given the great advances in computing technology, we can model proteins such as BPTI and their associated water molecules on a desktop computer in a couple of days, spending about 80 cents for electricity. Scientists have now simulated the aqueous ("in water") structures of more than 50 proteins and nucleic acids such as DNA.

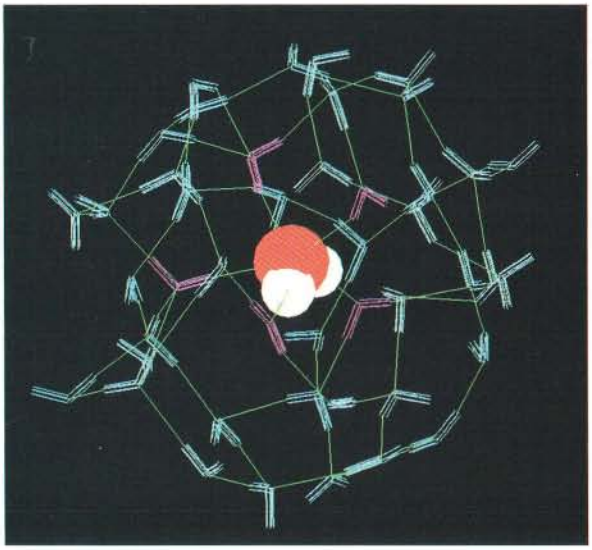
Why is it so important to understand the effects of water on the shapes of biological molecules? Principally, because a molecule's structure yields clues to how it functions, helping scientists decipher the biochemical interactions that add up to life. On a more practical level, understanding the structures of biological molecules in water may one day help researchers design new drugs that act by blocking or enhancing various biochemical pathways.

The Water Within

To understand how water affects the structures of biological molecules, we must first appreciate the distinctive properties of water itself. These properties stem from the unique structure of water and the way this structure allows water to "manage" the electric charges of other molecules.

A single water molecule (H_2O) has an essentially tetrahedral geometry, with an oxygen atom at the center of the tetrahedron, hydrogen atoms at two of the four corners and clouds of negative charge at the other two corners. The clouds of negative charge result from the way in which the atomic structures of oxygen and hydrogen combine. In simplified terms, oxygen has eight negatively charged electrons circling its positively charged nucleus: two in an inner shell and six in an outer shell. The inner shell's maximum capacity is two electrons, so it is full, but the outer shell can hold as many as eight. Hydrogen has

ALL ILLUSTRATIONS BY MARK GERSTEIN AND MICHAEL LEVITT



HYDROGEN BONDS give water its unique properties. In this model of liquid water, the central molecule (*red and white spheres*) has formed hydrogen bonds (*green lines*) with five other water molecules (*pink Vs*). Its hydrogen atoms (*white*) have bonded with the oxygens of two other waters, and its oxygen atom (*red*) has bonded with a hydrogen from each of three other water molecules. Each molecule of liquid water usually forms four or five hydrogen bonds.

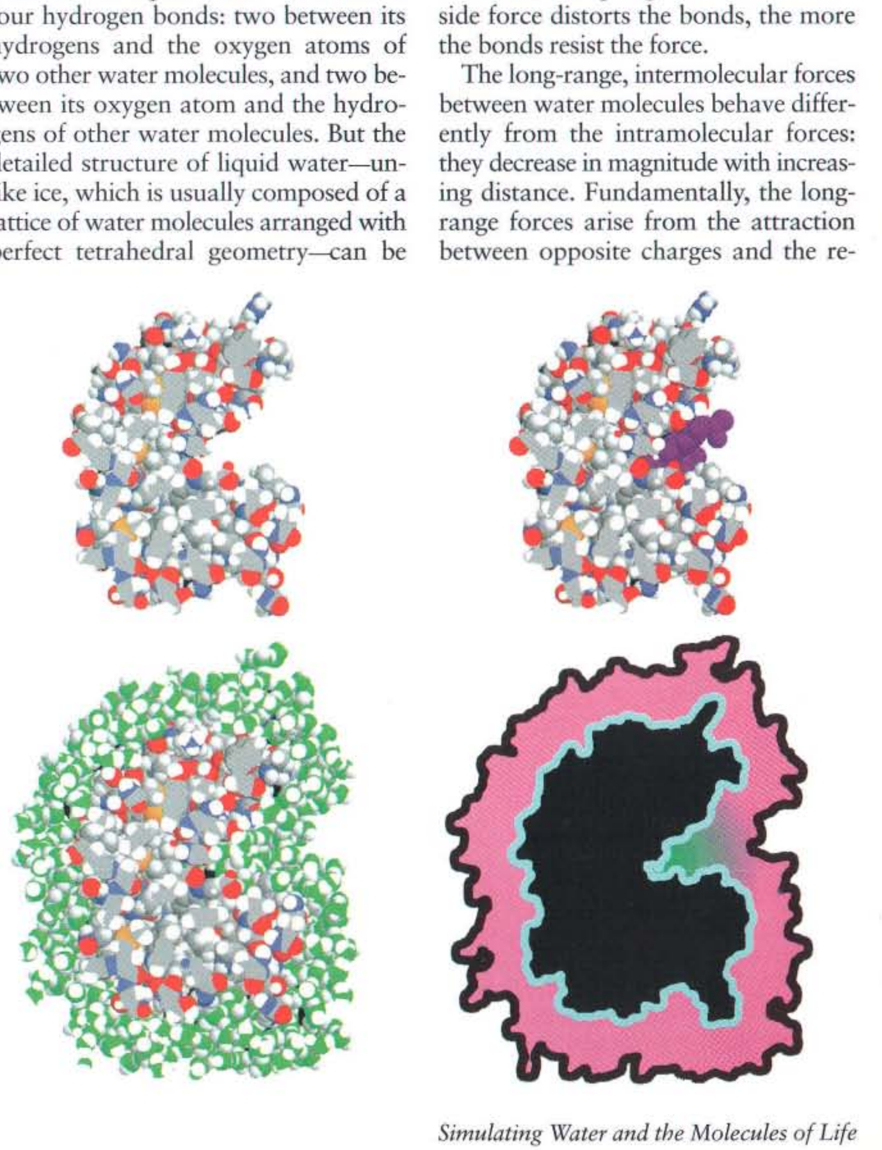
only one electron. When oxygen combines with two hydrogens, it attracts each hydrogen's electron in an attempt to fill its outer shell. Because each hydrogen electron spends more time around the oxygen atom than around its own positively charged nucleus, a water molecule is polar: it has two clouds of slight negative charge around the oxygen atom, and its two hydrogen atoms are left with slightly positive charges. These two types of charges counterbalance one another, however, so that water molecules are electrically neutral.

Chemists usually do not draw the clouds of negative charge around the oxygen atom of a water molecule; instead they conventionally depict water in a V shape [*see illustration above*]. Each side of the V corresponds to an oxygen-hydrogen bond roughly 10^{-8} centimeter in length. The angle formed between the two sides of the V is close to 105 degrees—slightly less than the 109.5-degree angle formed between any two sides of a perfect tetrahedron.

Because of the polarity of water molecules, interactions between a positively charged hydrogen of one water molecule and the negatively charged oxygen of another are favorable. These interactions are called hydrogen bonds. Reflecting water's tetrahedral geometry, each molecule in liquid water often forms four hydrogen bonds: two between its hydrogens and the oxygen atoms of two other water molecules, and two between its oxygen atom and the hydrogens of other water molecules. But the detailed structure of liquid water—unlike ice, which is usually composed of a lattice of water molecules arranged with perfect tetrahedral geometry—can be quite random and irregular. The actual number of hydrogen bonds per liquid water molecule ranges from three to six, with an average of about 4.5. The necessity of maintaining a tetrahedral, hydrogen-bonded structure gives water an "open," loosely packed structure compared with that of most other liquids, such as oils or liquid nitrogen.

To construct a computer model of water, we need to take into account two different types of forces: intramolecular and intermolecular. The interactions *within* a water molecule are modeled in terms of the short-range, springlike forces created by the chemical bonds between each molecule's hydrogens and oxygen. The interactions *between* water molecules are modeled in terms of long-range, electrical forces. The intramolecular forces restrain the lengths of the bonds between the oxygen of each water molecule and its hydrogens—and the angle formed between each of these bonds—to certain set values. These forces behave like springs: the more an outside force distorts the bonds, the more the bonds resist the force.

The long-range, intermolecular forces between water molecules behave differently from the intramolecular forces: they decrease in magnitude with increasing distance. Fundamentally, the long-range forces arise from the attraction between opposite charges and the re-



ACTIVE SITE of lysozyme—a naturally occurring enzyme that kills bacteria by breaking down sugary molecules in their cell walls—is in the protein's main groove (*upper left*). The groove is shaped precisely to accommodate the molecules it cleaves (*purple spheres, upper right*). Modeling how water (*green and white spheres, lower left*) interacts with the groove helps scientists to create a map of the active site (*lower right; green shading indicates easily displaced water molecules*). Such maps can be key in designing new drugs to block or enhance a particular enzyme's activity.

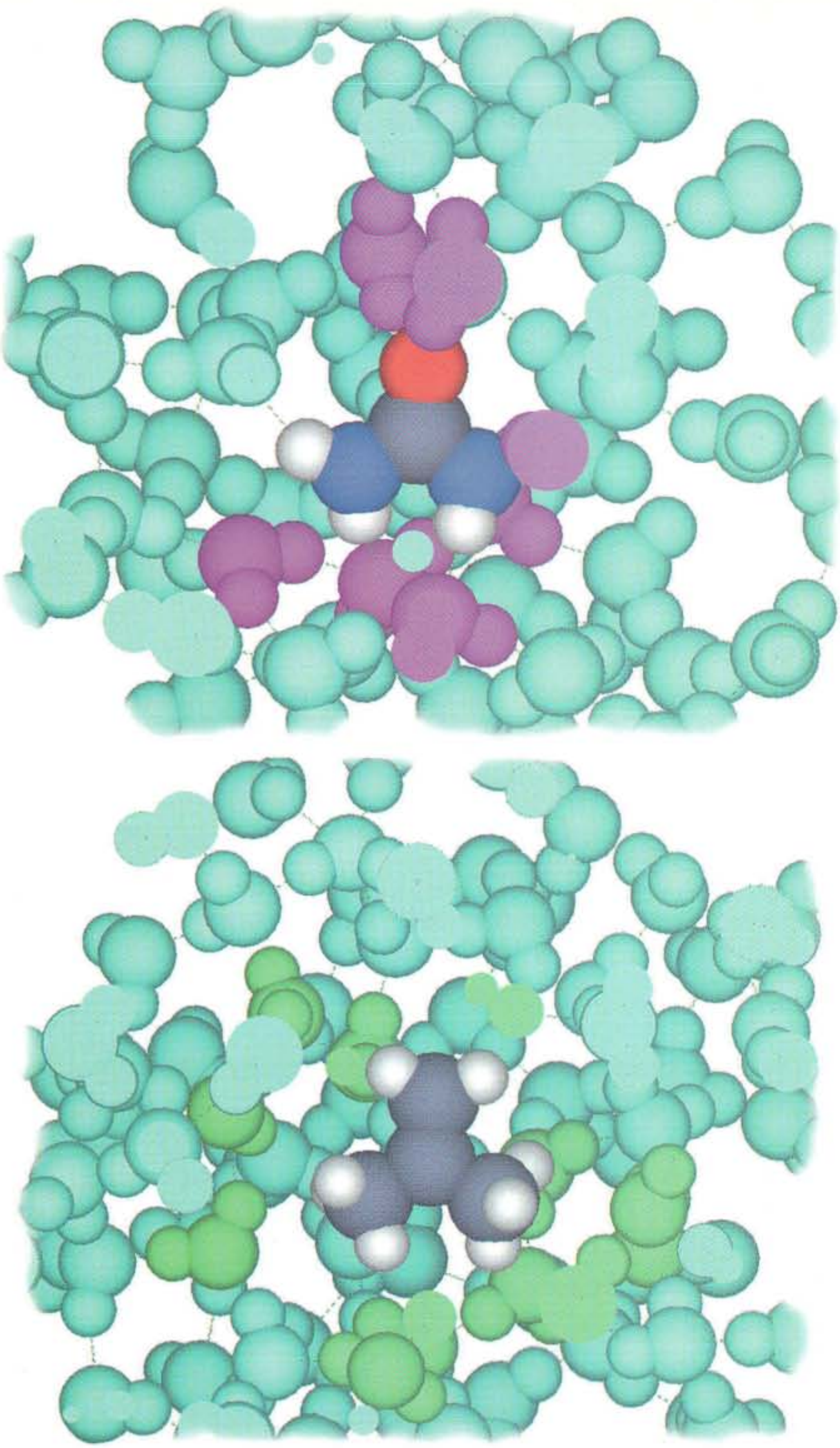
pulsion between similar charges. These forces give rise to hydrogen bonds as well as to weaker attractions called van der Waals forces.

The computer simulation of water molecules was pioneered in the late 1960s by Aneesur Rahman and Frank H. Stillinger of Bell Laboratories. Rahman and Stillinger simulated the motion of 216 water molecules in a rectangular box. (The researchers chose to model 216 water molecules because that number constitutes a box of waters six molecules deep, six molecules high and six molecules wide.) In their five-picosecond simulation—the longest possible using the computing technology of the time—Rahman and Stillinger found that the behavior of water is a direct consequence of the energetic relations among the water molecules. The simulation was able to reproduce quantitatively many of the bulk properties of water, such as its average structure, rate of diffusion and heat of vaporization.

Simulating Life

The importance of water in living processes derives not only from its ability to form hydrogen bonds with other water molecules but also from its capacity to interact with various types of biological molecules. Because of its polar nature, water readily interacts with other polar and charged molecules, such as acids, salts, sugars and the various regions of proteins and DNA. As a result of these interactions, water can dissolve polar molecules, which are consequently described as hydrophilic ("water-loving"). In contrast, water does not interact well with nonpolar molecules such as fats, giving rise to the observation that oil and water do not mix. Nonpolar molecules are therefore termed hydrophobic ("water-fearing").

Biological molecules such as proteins and DNA contain both hydrophilic and hydrophobic parts arranged in long chains. The three-dimensional structures of these molecules are dictated by the way these chains fold into more compact arrangements, so that hydrophilic groups are on the surface where they can interact with water, and hydrophobic groups are buried in the interior, away from water. In 1959 Walter Kauzmann proposed that such a hydrophobic effect was crucial for protein folding, and the role of hydrophobicity in protein folding is still a subject of great interest today [see "The Protein Folding Problem," by Frederic M. Richards; Scientific American, January 1991].



MOLECULES with nearly identical shapes interact differently with water depending on whether they are polar, or have partial charges on some of their atoms, or lack charges and are therefore nonpolar. Urea, a polar molecule found in urine, forms hydrogen bonds with water molecules (*purple spheres, top*). In contrast, the nonpolar isobutene does not form such bonds; instead water molecules hydrogen-bond with one another around isobutene, forming a cagelike structure (*green spheres, bottom*).

There are three types of waters that must be considered when building a computer model of a biological molecule in aqueous solution: the "ordered waters" that surround and strongly interact with the molecule, the "bulk waters" beyond and any waters that may be buried within the molecule. A single cell contains billions of water molecules. Almost all the space not occupied by the atoms of biological molecules is filled with water. Human cells are, in fact,

mostly water; the human body is roughly 60 percent water by weight.

How do we model all these waters, together with the individual atoms of a biological molecule? In simple terms, we first describe the basic interactions between all the atoms and then let the system evolve according to the laws of Newtonian physics. Such a simulation requires two basic ingredients: a way to describe the interactions within and among water and biological molecules—the intramolecular and intermolecular forces, in other words—and a procedure for charting their movements through time, called molecular dynamics.

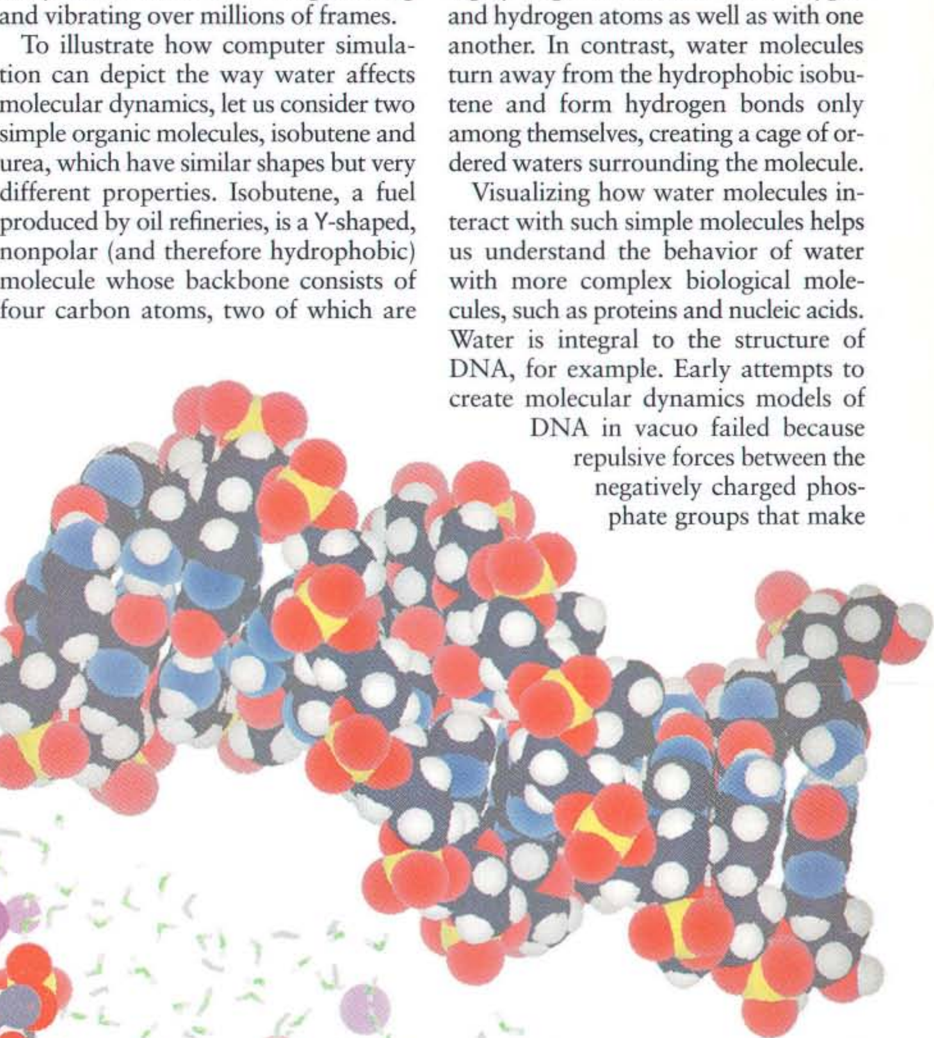
Molecular dynamics produces a sequence of configurations very much like frames in a movie. Each atom moves through time in a series of discrete steps, called timesteps. Essentially, the new position of an atom is its old position plus the distance it traveled during a given timestep. If no forces acted on an atom, the distance it traveled would be a function of its velocity at the previous position, because distance equals speed multiplied by time. During a timestep, however, the forces exerted by other atoms cause the atom to accelerate, which in turn changes its velocity. If the forces are constant during the timestep, Newton's laws dictate that the change in velocity is proportional to the force, so we can calculate an updated velocity. We then use this updated velocity to calculate the new position of the atom. The strongly interacting atoms of a liquid cannot move very far, and it is necessary to use a very short timestep: one femtosecond (10^{-15} second). During this period, a water molecule moves only $1/500$ of its diameter.

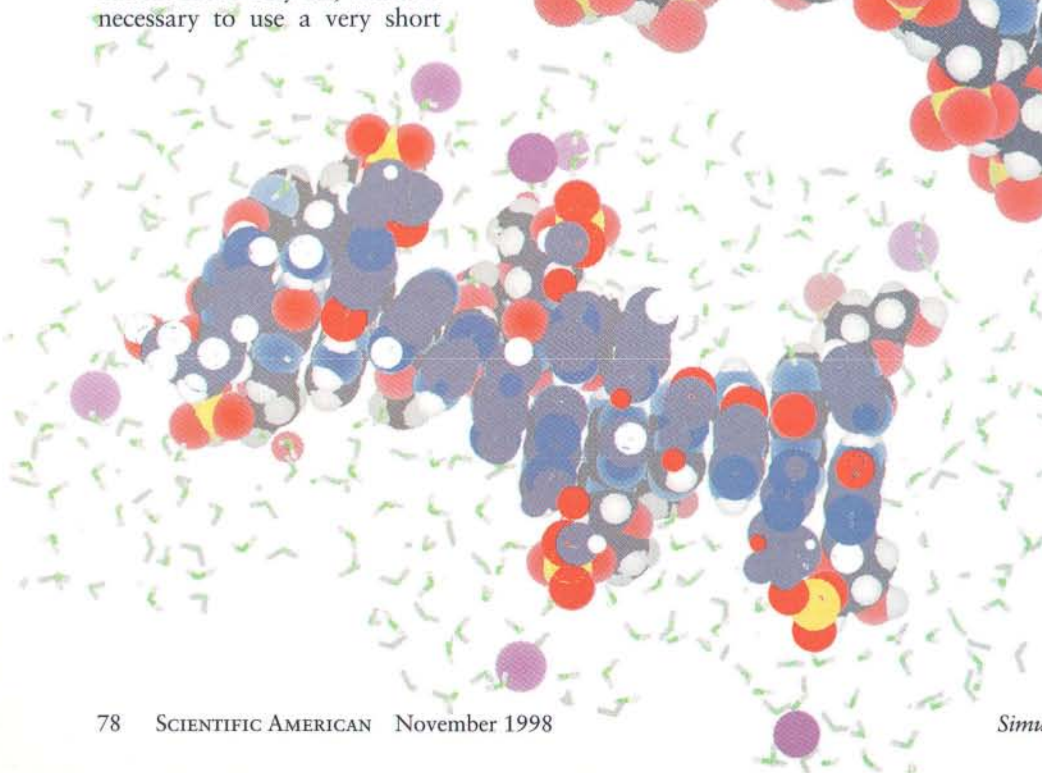
In a long simulation, calculating each timestep for all the atoms in a biological molecule with its ordered waters yields an enormous amount of data. A small protein in water, for instance, produces half a million sets of Cartesian coordinates in a nanosecond, each describing the positions of about 10,000 atoms. The movie generated by such a simulation is exquisitely detailed. We can see every water molecule rotating, shifting and vibrating over millions of frames.

To illustrate how computer simulation can depict the way water affects molecular dynamics, let us consider two simple organic molecules, isobutene and urea, which have similar shapes but very different properties. Isobutene, a fuel produced by oil refineries, is a Y-shaped, nonpolar (and therefore hydrophobic) molecule whose backbone consists of four carbon atoms, two of which are linked by a double bond. Urea is a product of protein metabolism that is excreted in urine. It, too, has a Y-shaped structure: a carbonyl group (C=O) linked to two amino groups (NH_2). Unlike isobutene, urea is a strongly polar molecule that is hydrophilic.

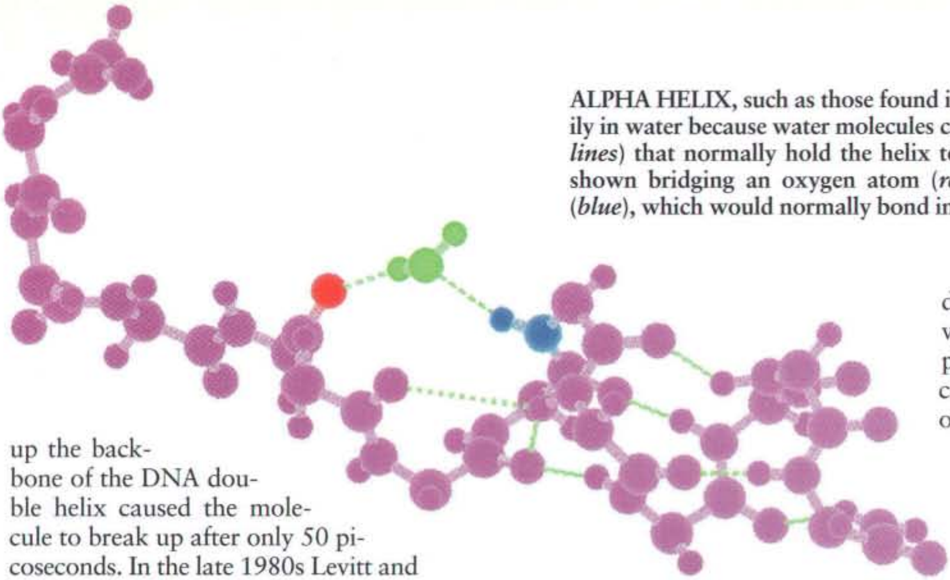
When we carry out molecular dynamics simulations of isobutene and urea, we see that water behaves differently around the two molecules [*see illustration on preceding page*]. Water molecules interact directly with urea, forming hydrogen bonds with urea's oxygen and hydrogen atoms as well as with one another. In contrast, water molecules turn away from the hydrophobic isobutene and form hydrogen bonds only among themselves, creating a cage of ordered waters surrounding the molecule.

Visualizing how water molecules interact with such simple molecules helps us understand the behavior of water with more complex biological molecules, such as proteins and nucleic acids. Water is integral to the structure of DNA, for example. Early attempts to create molecular dynamics models of DNA in vacuo failed because repulsive forces between the negatively charged phosphate groups that make





DNA DOUBLE HELIX (*top*) has two phosphate backbones (*red and yellow spheres*) wrapped around a stack of base pairs (*gray, blue, red and white spheres*). A slice through one frame of a DNA simulation in water (*bottom*) shows how water molecules (*green and white Vs*) can penetrate deeply into DNA's double-helical structure and stabilize it. Purple spheres indicate sodium ions in the water solution.



ALPHA HELIX, such as those found in most proteins, can unfold more easily in water because water molecules can replace the hydrogen bonds (*green lines*) that normally hold the helix together. A water molecule (*green*) is shown bridging an oxygen atom (*red*) and a nitrogen-hydrogen group (*blue*), which would normally bond in the folded helix.

up the backbone of the DNA double helix caused the molecule to break up after only 50 picoseconds. In the late 1980s Levitt and Miriam Hirshberg of the National Institute for Medical Research in London succeeded in making a 500-picosecond model of DNA by including water molecules that stabilized the double-helical structure by forming hydrogen bonds with the phosphate groups. Subsequent simulations of DNA in water have revealed that water molecules are able to interact with nearly every part of DNA's double helix, including the base pairs that constitute the genetic code.

In contrast, water is not able to penetrate deeply into the structures of proteins, whose hydrophobic regions are tucked on the inside into a close-fitting core. Accordingly, protein-water simulations have focused on the protein surface, which is much less tightly packed than the protein interior.

The way in which water molecules interact with the surfaces of proteins results in much interesting geometry—particularly in the deep grooves on the surfaces of enzymes, proteins that foster chemical reactions in cells. Hydrogen-bonded water molecules have difficulty fitting into these grooves and are easily displaced by ligands—the molecules with which an enzyme is intended to interact—which might explain why the active sites of enzymes frequently occur in grooves. We often find that the arrangement of water molecules in an empty active site mimics the geometry and structure of the actual ligand, knowledge that is sometimes used in drug design.

Living in the Real World

How closely do simulations of biological molecules in water resemble reality? Unfortunately, we cannot answer this question definitively, because no experimental technique can provide as much detailed information about individual molecules and their interactions as computer modeling can. What we can do is to compare various aggregated and averaged values derived from simulations with experimental results.

One of most important approaches that can be used to verify the structures of biological molecules simulated in water is neutron and x-ray scattering. In a neutron-scattering experiment, we direct a beam of neutrons at a small sample and record how the neutrons are scattered by the molecules that make up the sample. Each space between the molecules acts as a tiny slit, yielding a characteristic diffraction pattern. By analyzing these patterns, we can readily determine the spacing between the various molecules. When we compare neutron-scattering results with computer simulations, we find that, on average, these distances coincide.

To confirm the dynamics of a molecular simulation, we compare the predicted behavior of the simulated biological molecule in water with its observed properties in the laboratory. For example, most proteins contain at least one alpha helix, where the amino acids that make up the protein twist to form a coil. We know from experiments that heat causes these alpha helices to uncurl, yet in early attempts to simulate the behavior of a simple alpha helix in vacuo at elevated temperatures, the helix remained intact. Only by adding water to the simulation were Levitt and Valerie Daggett of the University of Washington able to mimic an alpha helix's actual behavior [*see illustration above*].

Such computer simulations are yielding more and more information about the shapes of various biological molecules and how they perform their jobs in a living organism. We are, however, constantly running up against the limitations of computing technology and the cost of supercomputer time as we seek to conduct simulations of increasingly complex biological molecules in their watery environments. When scientists publish models of biological molecules in journals, they usually draw their models in bright colors and place them against a plain, black background. We now know that the background in which these molecules exist—water—is just as important as they are. SA

The Authors

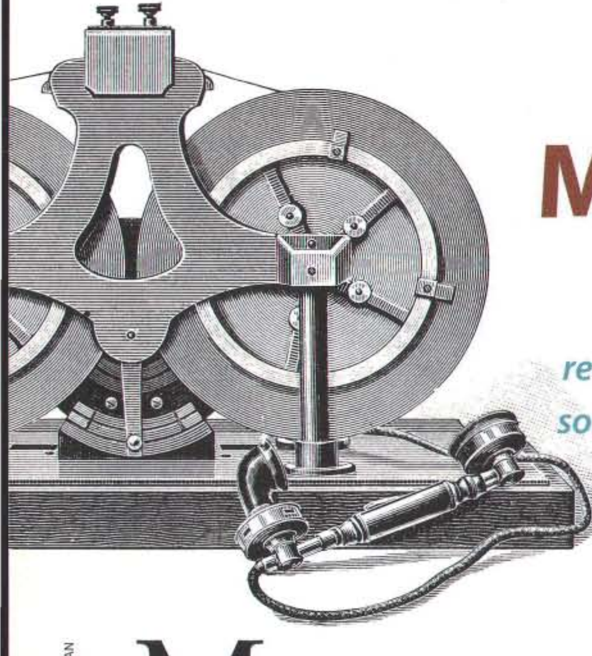
MARK GERSTEIN and MICHAEL LEVITT have collaborated since 1993, when Gerstein became a postdoctoral fellow in Stanford University's department of structural biology, which Levitt still chairs. Levitt obtained his Ph.D. in 1971 from the University of Cambridge. He has held academic positions at the Laboratory of Molecular Biology in Cambridge, England, the Salk Institute for Biological Studies in San Diego and the Weizmann Institute of Science in Rehovot, Israel. A frequent consultant for pharmaceutical companies, Levitt also founded the biotechnology company Molecular Applications Group in Palo Alto, Calif. Gerstein is an assistant professor at Yale University. He completed his Ph.D. at Cambridge in 1993.

Further Reading

WATER: NOW YOU SEE IT, NOW YOU DON'T. Michael Levitt and Britt H. Park in *Structure*, Vol. 1, No. 4, pages 223–226; December 15, 1993.

PACKING AT THE PROTEIN-WATER INTERFACE. Mark Gerstein and Cyrus Chothia in *Proceedings of the National Academy of Sciences USA*, Vol. 93, No. 19, pages 10167–10172; September 17, 1996.

For electronic archives of molecular structures, visit bioinfo.mbb.yale.edu and hyper.stanford.edu on the World Wide Web.



SCIENTIFIC AMERICAN

100 Years of Magnetic Memories

Although the technology is ubiquitous today, magnetic recording had a sluggish start. The underlying science was something of a mystery, applications were slow to emerge, and business and politics stifled development

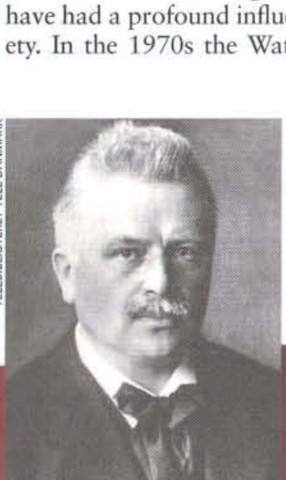
by James D. Livingston

Magnets store much of the world's information: data on computer disks, entertainment on video and audiotapes, messages on telephone answering machines and account information on the coated stripes of ATM and credit cards. All these different media preserve words, numbers, images and sounds as invisible patterns of north and south poles. The technology is magnetic recording, which celebrates its centennial this year.

In recent decades magnetic memories have had a profound influence on society. In the 1970s the Watergate tapes from the Oval Office provided the "smoking gun" that forced President Richard Nixon to resign. This year audiotapes that disclosed an improper relationship between President Bill Clinton and a former White House intern led to another executive scandal.

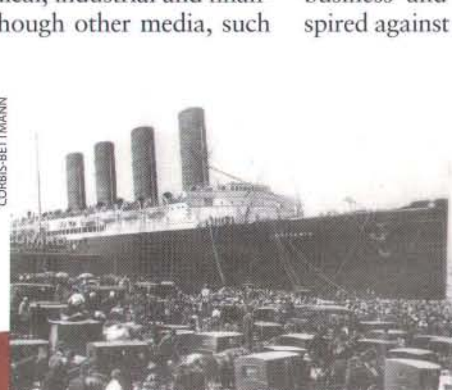
With less notoriety, the magnetic memories of computers have been quietly keeping track of people, goods and money for many years. In the information age, they have served as the warehouses of electronic knowledge, storing scientific, medical, industrial and financial data. Although other media, such as film, paper and CD-ROMs, are also widely used, magnetic recording boasts one huge advantage that distinguishes it from most competing technologies: it can easily rerecord or erase information (as demonstrated by secretary Rose Mary Woods and the infamous gap in the Nixon tapes).

But magnetic recording, introduced 100 years ago, was an invention that languished for decades. Some of the underlying physics was unknown, suitable applications were not quite ready, and business and political obstacles conspired against early adoption. In its sec-



TELEBIBLIOTEKET-TELE DANMARK

1898
VALDEMAR POULSEN, a Danish engineer, invents the telegraphone (*above*), a device that records the human voice magnetically on steel wires or ribbons.



CORBIS-BETTMANN

1915
LUSITANIA is sunk by a U-boat, heightening U.S. suspicions that the Germans are using telegraphones to record messages for high-speed transmissions.

UPI/CORBIS-BETTMANN

1900
FRANZ JOSEPH, Austrian emperor, records his voice with a telegraphone at the Paris Exposition, where the device is a sensation.

EMTEC MAGNETICS

1928
FRITZ PFLEUMER, an Austrian chemist, invents a device that records on a lightweight tape coated with magnetic particles.

ond half-century, however, the technology quickly became an invaluable tool of modern society.

The First Magnetic Memories

People have known for some time that certain materials "remember" the direction in which they were magnetized. As early as 200 B.C., Chinese magicians of the emperor's court made compasses out of lodestones, iron-rich rocks that are strongly magnetic.

But it was not until the 1800s, when scientists began to uncover the link between electricity and magnetism, that magnets began to play a key role in such notable inventions as telegraphs, telephones, generators, transformers and motors. Then, near the end of the 19th century, a magnetic material was used to record and reproduce the human voice. The device, called a telegraphone, was patented in 1898 by Valdemar Poulsen, a Danish engineer.

An employee of the Copenhagen Telephone Company, Poulsen thought that people would want a device to record telephone messages. He demonstrated the concept of his invention to his friends with a steel piano wire stretched across his laboratory. Sliding an electromagnet along the wire, Poulsen began yelling into a telephone mouthpiece connected to the magnet. The device converted his words into an electric signal, which was then fed into the electromagnet to generate a magnetic field that changed with the volume, pitch and other characteristics of his voice. The varying field was then imprinted along the steel wire.

After Poulsen reached the end of the wire, he returned to the starting point and replaced the mouthpiece with a receiver. When one of his friends slid the electromagnet along the wire, the device worked in reverse, first detecting the field in the wire, converting that time-varying information into an electric signal (via electromagnetic induction) and turning the current back into sound. Poulsen's friend could hear the inventor's voice faintly in the receiver!

Poulsen quickly refined his invention—in one variation, the steel wire was wrapped around a cylinder—and applied for patents in various countries. The initial response, though, was uniformly negative. In the U.S. a patent examiner wrote that the telegraphone would fail because its claims were "contrary to all known laws of magnetism." Apparently, the laws of magnetism were not *all* known in those days, because Poulsen's invention did indeed work.

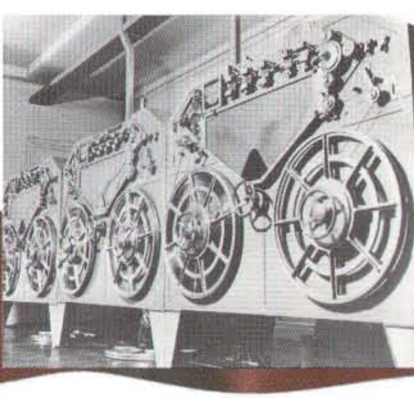
A series of demonstrations in Europe soon proved the device. At the Paris Exposition in 1900, the telegraphone was a spectacular success, impressing many of the visitors, among them Emperor Franz Joseph of Austria, who recorded a message that remains the oldest magnetic recording in existence today.

Of course, Poulsen was not the first to record and reproduce the human voice. More than two decades earlier, Thomas Edison had patented the phonograph, which originally recorded sounds with grooves cut in tin foil wrapped on a cylinder. Later, wax cylinders and then plastic disks were used, and by the time Poulsen's invention came to the world's attention, the phonograph was already an established product.

Even so, the telegraphone appeared to have a technical edge. An article in *Scientific American* in 1900 stated that the sound produced by the telegraphone "is very distinct and is entirely free from the disagreeable scratching noises generally heard in the phonograph." Nevertheless, another half century would pass before magnetic recording found widespread application.

A Slow Start

What stalled the development of such an obviously promising invention? Historians blame a combination of business and technological factors. Initially, the American Telephone & Telegraph Company (AT&T) resisted telegraphones, estimating that the company would lose up to a third of its business if callers felt their conversations might be recorded. In 1906 the magazine *Technical World* carried an article



AMPEX ARCHIVES

1930s
BLATTNERPHONE type of magnetic recorder is used by the BBC for rebroadcasting radio programs. Inserting the reels of heavy steel ribbon requires two people.

1933
ADOLF HITLER becomes chancellor of Germany. The Gestapo purchases large quantities of magnetic recorders for interrogations.

UPI/CORBIS-BETTMANN

1936
LONDON PHILHARMONIC ORCHESTRA is recorded in concert at Ludwigshafen, Germany, by a Magnetophon, which uses plastic tape coated with iron oxide. The recording survives today.

EMTEC MAGNETICS

entitled "A Spool of Wire Speaks," which opened with a scene between two imaginary businessmen, Jones and Brown. The men were arguing heatedly over what Brown had—or had not—said in a previous telephone conversation. With no agreement in sight, Jones reached into his desk and withdrew a spool of steel wire, which he then inserted into a telegraphone. The recording established, to Brown's dismay, that Jones was right.

Because of such privacy concerns, the first U.S. application of magnetic recording was in dictating machines, which were manufactured by the American Telegraphone Company, founded in 1903. But competing phonographic devices such as Ediphones and Dictaphones had the advantage of a 20-year head start: they were cheaper, easier to use and more reliable.

Unfortunately, refinement of the telegraphone was hampered by a meager understanding of the basic physical phenomenon on which the device depended. The conversion of sound into electricity and then into a magnetic field was fairly well understood from earlier experience with the telephone. But the mechanism by which the history of those magnetic fields could be stored in a recording medium, such as a steel wire, was something of a mystery at the time. Another limitation was the low level of playback sound.

Perhaps American Telegraphone's biggest problems, though, were not technical. Among the scant number of telegraphones sold, several were installed in transatlantic wireless stations in Tuckerton, N.J., and Sayville, N.Y., which were then operated, respectively, by the German company Telefunken and by a U.S. firm connected to German interests. It was also known that the German navy had purchased telegraphones for its submarines.

With the outbreak of World War I, the two East Coast wireless stations were suspected of transmitting military information to German submarines in the Atlantic. One allegation was that telegraphones were being used to record secret messages that could then be sent as high-speed wireless signals. These signals could be recorded by a telegraphone on a submarine, where the transmission could be played back at normal speed. Suspicions heightened when a radio ham in New Jersey discovered that the mysterious buzzing he had been picking up at night evolved into a series of Morse-code dots and dashes when it was played at a slower speed.

The U.S. government seized the Tuckerton station in 1914, and after a U-boat sank the British liner *Lusitania* off the coast of Ireland, the Sayville station was likewise confiscated. The U.S. Navy itself had purchased 14 telegraphones, but the devices failed to perform properly. Later, accusations of treason were brought against officers of American Telegraphone, which had already been plagued by bad management and paltry sales (only a few more than 100 dictating machines had been sold). Not surprisingly, the ill-fated company was soon in receivership, and the commercial development of magnetic recording in the U.S. would essentially disappear for two decades.

A Smoother Path in Europe

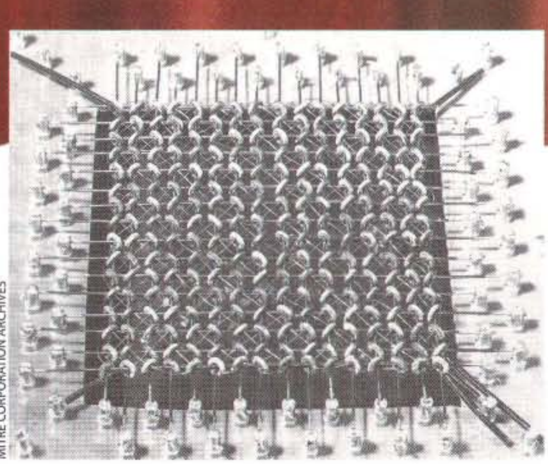
Magnetic recording fared much better in Europe. By the 1920s, electronic amplifiers had solved the low-volume problem, and German inventor Kurt Stille was working on modified telegraphones with improved recording capacity. As a climax to his sales talks, Stille (whose name in German ironically means "silence") would reach into his pocket and pull out a steel cylinder about 20 centimeters (eight inches) long. "On this little tube," he would declare, "I can record a whole symphony!"

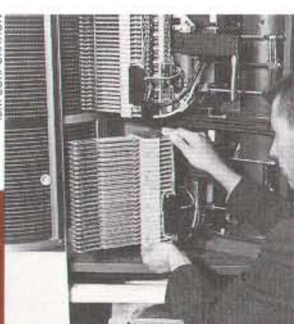
One licensee of Stille's technology was the Ludwig Blattner Picture Company in Britain, which developed a magnetic recorder called the Blattnerphone. Originally aimed at the production of sound films, the device instead found its niche in the rapidly developing radio business. Starting in 1931 the British Broadcasting Company (BBC) used the Blattnerphone for worldwide rebroadcasts of radio programs, for immediate playback at rehearsals and for preserving important speeches and news events.

1947
BING CROSBY's radio show becomes the first major U.S. program to be aired from magnetic recordings.


UPI/CORBIS-BETTMANN

1952
CORE MEMORY for Whirlwind I computer stores 256 bits of data.


MITRE CORPORATION ARCHIVES


IBM CORPORATION

1950s
HARD-DISK technology is developed for storing computer data magnetically, and IBM ships the first such device in 1957. A later product (*above*) is equipped with 50 disks for a total capacity of 56 megabytes.

The Blattnerphone was a formidable machine—roughly one and a half meters (five feet) high, one and a half meters wide and more than a half meter deep. And it had a mass of about 900 kilograms (nearly a ton). The recording medium was a steel ribbon, approximately three millimeters (an eighth of an inch) wide, which moved past the heads at a speed of about a meter or so per second. A half-hour recording required roughly one and a half kilometers (about a mile) of ribbon.

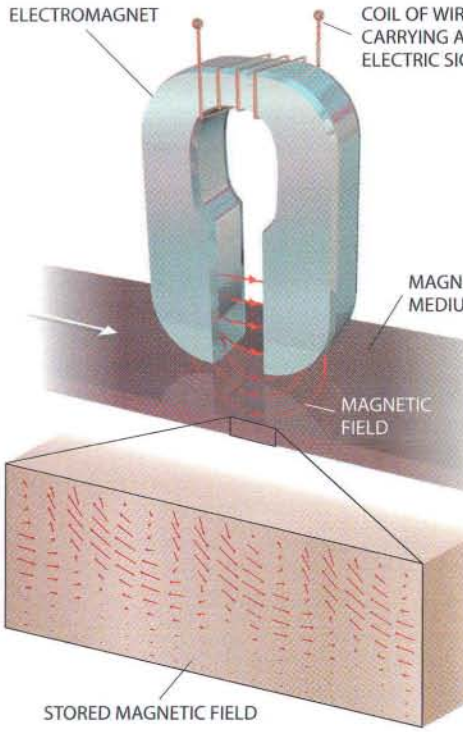
Still, the Blattnerphone had its advantages. Whereas editing with phonographic disks required rerecording, technicians could cut and solder the steel ribbons with relative ease. Unfortunately, the soldered joints occasionally snapped, resulting in dangerous bands of flying steel. But the machines generally worked well, and variations of the Blattnerphone remained in operation in several countries as recently as 1945.

Back in Germany in 1925, Stille and his colleagues had manufactured and marketed a device that was a combination dictating machine and telephone recorder. In contrast with AT&T, several European telephone companies had allowed magnetic recorders to be connected to their lines. The market for dictating machines in Germany, however, was very limited because of the difficult financial times there, which made stenographers a cheap labor source.

An improved version of Stille's machine was marketed by the Lorenz Company in 1933, around the time that Adolf Hitler came to power, and the Gestapo was soon buying huge quantities of the products to record interrogations and telephone conversations. Lorenz also developed a steel-ribbon recorder that was adopted by the German Broadcasting Company, which, along with the BBC and others, made magnetically recorded radio programs common in Europe in the 1930s. Across the Atlantic, in the U.S., network radio continued to restrict itself almost exclusively to live programming.

Despite their success in Europe, the early magnetic recorders had one severe drawback: they required extravagant lengths of heavy steel in the form of wires or ribbons. Scientists realize today that solid steel was a rather poor choice for a recording medium. Thus, a crucial step in the evolution of the technology occurred in 1927, when Austrian chemist Fritz Pfleumer developed a paper tape coated with powdered magnetic particles. Actually, such an invention had been among the possibilities suggested in an 1888 magazine article by Oberlin Smith, an American engineer. In that publication, Smith had also proposed a magnetic recorder, but his ideas never reached fruition, leaving Poulsen to become the first to produce such a device.

Pfleumer had been developing a technique for manufacturing gold-colored paper for cigarette tips by dispersing small particles of bronze in glue and applying the mixture to the paper's sur-



BRYAN CHRISTIE

MAGNETIC RECORDING can store audio, video or computer data. The information is first converted into an electric signal, which is then transmitted through the coiled wire of an electromagnet. The resulting time-varying magnetic field (*red*) is then imprinted in a magnetic medium that is moving relative to the electromagnet. If the electric signal is a sine wave, the resulting pattern in the magnetic medium will also be somewhat sinusoidal in nature.

AMPEX ARCHIVES

1956
VR-1000, the first commercial videotape recorder, is demonstrated in prototype by Ampex. The large, complex equipment uses tape that is five centimeters (two inches) wide.

1959
NIKITA KHRUSHCHEV, Soviet premier, and Richard Nixon argue while visiting an exhibit of home appliances in Moscow. The impromptu incident, recorded on videotape, becomes known as the Kitchen Debate.

NATIONAL ARCHIVES AND RECORDS ADMINISTRATION

1963
COMPACT AUDIOCASSETTE is introduced by Philips. The eight-track and microcassette formats will follow.

ARCHIVE PHOTOS (*audiocassette*); BLANK ARCHIVES/ ARCHIVE PHOTOS (*eight-track cartridge*)

face. Pfleumer realized he could use a similar process to coat a paper strip with steel particles, and in 1928 he built a magnetic recorder that used such a coated tape as its recording medium. The device boasted a simplified mechanism for moving the lightweight tape across a new type of magnetic head. Pfleumer's machine was so promising that by 1932 the electric company AEG had purchased all his patents.

AEG collaborated with the chemical company I. G. Farben (today BASF) to investigate the use of other magnetic particles on various types of tapes. They found that the best results could be obtained from a form of iron oxide coated on plastic tape. The research led to the invention of the "Magnetophon," which was first displayed to the public in 1935 at the Berlin Radio Exhibition. In November of the following year, the device taped the London Philharmonic Orchestra in concert at Ludwigshafen, Germany, in a recording that survives today. With the subsequent addition of AC biasing, which improves sound quality by superimposing a high-frequency electric signal on the recorded information, the Magnetophon soon surpassed all rivals. The device is considered the direct ancestor of the modern tape recorder.

Such overseas successes finally prodded AT&T to commence research on magnetic recording. Although this work did not lead to any immediate commercial products, it did enable the company to produce magnetic recorders for the U.S. military in World War II. AT&T was joined by other U.S. companies, such as Brush Development Corporation, under a project led by Semi Begun, an engineer who had worked with Stille before leaving Nazi Germany in 1935; the Armour Research Foundation, in work headed by Marvin Camras, an American engineer; and General Electric, which was a licensee of Armour. Together the companies manufactured thousands of magnetic wire recorders for the U.S. military to keep logs and store messages on planes and ships, as well as in the field.

But British and American experts monitoring Nazi broadcasts were aware that the Germans had a superior device. It was the Magnetophon, which the Allies discovered as they swept through Europe during the final months of the war. When U.S. troops stormed the studios of Radio Luxembourg, they reportedly found a Magnetophon delivering one of Hitler's last harangues from an advanced version of Pfleumer's coated tape. American servicemen dismantled several of the confiscated devices to ship them back to the U.S., where they would play a major role in stimulating postwar interest in magnetic recording.

One of those servicemen was Jack Mullin, who demonstrated a Magnetophon to crooner Bing Crosby in 1947. The popular entertainer had been doing his radio program live with NBC, but he disliked the constraints of live broadcasts. Because NBC then had a restriction against recorded shows, Crosby joined ABC, which had a more liberal policy. But the quality of the existing technology, which required recording and rerecording on 33 $1/3$-rpm phonographic disks, was often poor. After Crosby heard the Magnetophon, he was sold. His show became the first major radio program in the U.S. to be aired from magnetic recordings.

The U.S. Catches Up

For months, Mullin was using the machines and tapes he had shipped from Germany. But by 1948 Ampex Corporation had produced an improved version of the Magnetophon, and other manufacturers were soon marketing competing products. Meanwhile companies like 3M had developed better tapes. Within a few years, tape had displaced wire as the dominant magnetic recording medium, and taped radio programs had become commonplace in the U.S. The ease of editing magnetic tape facilitated various innovations, including the splicing of two or three takes of a song to arrive at an acceptable performance, the addition of canned audience laughter and the deleting of short segments of undesired material, such as Crosby's unintentional cough after say-


UPI/CORBIS-BETTMANN

1974
RICHARD NIXON resigns from the presidency after audiotapes from the Oval Office reveal his role in the Watergate scandal.

KTLA-CHANNEL 5

1991
POLICE BRUTALITY becomes public after videotape of Rodney King's beating is aired.


DONAL PHILBY *FPG International*

1971
FLOPPY DISK is introduced, first in eight-inch format and later in 5.25 inches (*left*) and 3.5 inches.

1998
MONICA LEWINSKY becomes the center of a lurid presidential scandal, many details of which are disclosed in secret telephone tapings.


KEN CEDENO *Sipa*

ing, "If you like smoking," in a commercial for Chesterfield cigarettes.

Camras and others also promoted magnetic recording for the sound tracks of motion pictures, which since the birth of the "talkies" had used an optical system that converted sound into a light signal of varying intensity, which could then be captured on film. Because magnetic recording greatly simplified the editing and mixing of dialogue, music and sound effects, the technology caught on quickly. By 1951 the overwhelming majority of Hollywood's original production recording was done magnetically, even though the actual sound tracks delivered to theaters were still mainly optical.

Another entertainment medium that quickly embraced magnetic sound recording in the postwar years was television, which by 1949 was already using more film than Hollywood was. And on the horizon was a related technology that promised to be ideal for TV: magnetic *video* recording. First, however, engineers had to overcome a major hurdle.

Audio recording requires the reproduction of signals with frequencies up to 20,000 cycles per second—the uppermost limit detectable by the human ear. Video signals, on the other hand, must carry and deliver considerably more data. To produce the illusion of continuous motion, TVs in the U.S. display 30 complete images per second (in much of Europe and many other places, the number is 25), each consisting of several hundred horizontal lines with every line containing hundreds of fine dots.

Instead of trying to accelerate the tape speed, video engineers cleverly mounted the record/playback head on a rotating wheel to increase the *relative* motion between the head and the moving tape. Using such a mechanism, Ampex introduced the first commercial video recorder in 1956. Today taped TV broadcasts are so commonplace that the few shows that are not recorded often boast that they are live.

World War II's influence in magnetic recording reached far beyond the entertainment field. The U.S. war effort also accelerated the development of electronic digital computers. One of the first such machines—Whirlwind I, developed at the Massachusetts Institute of Technology in the 1940s and 1950s—had a "core memory," a two-dimensional array of tiny doughnut-shaped ferrite magnets. The toroids could be magnetized either clockwise or counterclockwise to represent a 0 or 1 in binary. Other forms of magnetic computer memories have included tape (UNIVAC I, the first commercial computer in the U.S., relied on magnetic tape for its auxiliary memory), drums (coated cylinders with tracks around their circumference) and disks (both hard and floppy).

In the 1970s integrated circuits replaced core memories, but magnetic hard disks are still the primary form of data storage inside a computer. More recently, industry has made rapid advances in increasing the storage densities on disks. Surface layers of iron-oxide particles have given way to thin films of cobalt-rich alloys, which are more magnetic than the oxides. In addition, the technology of the read heads is migrating from electromagnetic induction to magnetoresistive, in which the varying magnetic field in a disk is detected as changes in the electrical resistance of the overlying read head [see "Data-Storage Technologies for Advanced Computing," by Mark H. Kryder, SCIENTIFIC AMERICAN, October 1987]. Consequently, despite growing competition from laser-based optical systems such as CD-ROMs, magnetic recording in all its forms (including a hybrid approach called magneto-optical) remains the dominant technology for storing electronic information.

From the 1950s the public became increasingly aware of magnetic recording through continued rapid progress in consumer products. In 1963 Philips introduced the compact audiocassette, and within two decades the sales of such products surpassed those of phonograph records. Home-use videocassette recorders became available in the 1970s; many movies now earn more money through videotape rentals than they do in the theaters. Sales of camcorders, which were introduced in the 1980s, and the use of magnetically coded data on credit and ATM cards, hotel room keys and employee ID and access cards have also grown dramatically. And today many households have telephone answering machines—the very application that motivated Poulsen's invention a century ago.

To Tape or Not to Tape

But magnetic recording—with its ability to store human words and actions easily and surreptitiously—has raised ethical questions. For good or ill, the memory of an assembly of magnetic particles is far more objective and reliable than the subjective and often faulty electrochemical mechanism of the human brain.

The ethics of taping must balance the search for truth with the right to privacy. Although some people may question the ethics of tapings in which some or all of the participants are unaware that they are being recorded, the video images of Rodney King's beating raised awareness of police brutality in the U.S. In less celebrated cases, surveillance cameras have helped reveal abuses in certain nursing homes and hospitals. Video monitoring equipment in banks and court-approved wiretapping routinely bring many criminals to justice, and cockpit black-box recorders help to clarify the causes of plane crashes.

On balance, Poulsen's 1898 invention has surely brought more good than harm. The ability of magnetic materials to store data, sights and sounds has indeed become an invaluable technology of modern society. A century ago Poulsen could hardly have imagined the far-reaching consequences of his fledgling device. Today people might need a magnetic recorder to remember them all.

The Author

JAMES D. LIVINGSTON is a senior lecturer in the department of materials science and engineering at the Massachusetts Institute of Technology. He is the author of *Driving Force: The Natural Magic of Magnets* (Harvard University Press, 1996) and *Electronic Properties of Engineering Materials* (John Wiley & Sons, December 1998). Livingston received his Ph.D. in applied physics from Harvard University and is a member of the National Academy of Engineering and a fellow of the American Physical Society and of ASM International.

Further Reading

MAGNETIC RECORDING. S. J. Begun. Murray Hill Books, 1949.

MAGNETIC RECORDING HANDBOOK. Marvin Camras. Van Nostrand Reinhold, 1988.

MAGNETIC RECORDING: THE FIRST 100 YEARS. Edited by Eric D. Daniel, C. Denis Mee and Mark H. Clark. IEEE Press, 1999.

by Shawn Carlson

Floating a Challenge

Rarely does an unexpected call from a federal agency make my day; typically it just signals trouble. But I was pleasantly surprised by one recent phone call from Washington, D.C., which informed me of a fascinating project to chart ocean currents. My source, a scientist from the U.S. Geological Survey, explained that the program involves free-floating buoys (called drifters) that transmit their position along with various measurements of the ocean's properties to land-bound investigators. Government researchers would also like to study U.S. coastal waters in this way, yet they don't want to cause pollution and can't afford to deploy an armada of expensive drifters. The coordinator of this program asked me whether amateurs, whose skills have been honed by tight budgets, might have some clever ideas.

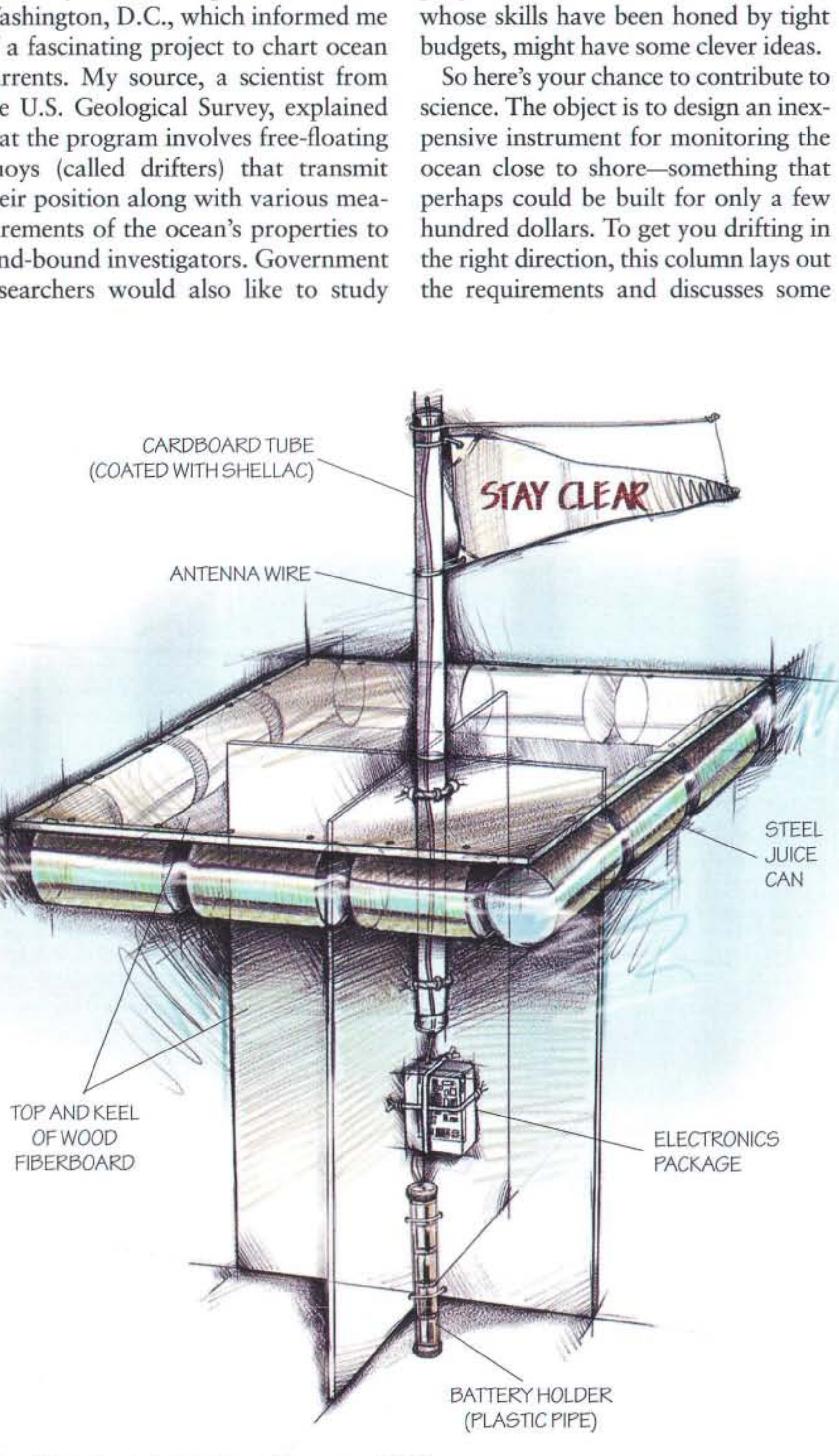
So here's your chance to contribute to science. The object is to design an inexpensive instrument for monitoring the ocean close to shore—something that perhaps could be built for only a few hundred dollars. To get you drifting in the right direction, this column lays out the requirements and discusses some hardware that may help. But don't limit yourself to the ideas presented here. I'll forward the best solutions to the federal agencies involved and post them on the World Wide Web site of the Society for Amateur Scientists. The winning designs in this informal competition will also earn laurels by appearing in a forthcoming column.

The drifter should be able to follow coastal currents for a full month, reporting its position and the condition of the sea around it. But it should be neither a danger to navigation nor a threat to marine life. The device should thus be free of items that might ensnare, choke, poison or otherwise injure marine organisms. The ideal drifter would be largely biodegradable. Or it could be made so that whatever parts don't break down quickly are encapsulated and sink promptly to the bottom, where they cannot do any harm.

Indeed, part of the challenge is first to decide just what poses an environmental danger. Only then can you design a drifter to be as benign as possible. Keep in mind also that the instrument must be small enough not to create a hazard to shipping and yet should be visible to swimmers, surfers and boaters so that they can steer clear. Finally, it should be able to survive storms or choppy seas, and it must move with the currents, not with the winds.

Finding suitably biodegradable parts will require some experimentation. Cardboard supports, for example, might rapidly become waterlogged. And floats made from steel cans may corrode too quickly. To slow these processes, you might coat such components with a layer of shellac, wax or some other natural substance.

My own conception is illustrated at the left. A large keel and a low profile above the water prevent the drifter from tipping and ensure that it moves with the currents. To warn off boaters, a colorful cotton flag attaches to the card-



DANIELS & DANIELS

***SIMPLE OCEANOGRAPHIC DRIFTER** designed along these lines, for example, could be constructed inexpensively.*

board tube that protects a radio antenna.

But my proposal merely shows one sample design. There are probably myriad ways to configure a drifter with the necessary attributes, and I look forward to judging your detailed plans. Because the drifter must be able to sense and report its position and the state of the sea around it, I will also select a separate winner for excellence in designing an electronics package that reliably accomplishes all these tasks for the least amount of money.

Though difficult, that challenge is easier than it might at first seem. Satellites of the Global Positioning System (GPS) make it child's play to determine location. Garmin, for instance, sells tiny GPS receivers for less than $200 that could locate a drifter anywhere on the earth to within 100 meters or so. (See www.garmin.com/oemTracPaks.html or call 913-397-8200.)

Telemetry is, perhaps, trickier. Drifters deployed in the open ocean use the ARGOS satellites to transmit their data home. But at least for the moment, this service is quite pricey. Fortunately, for coastal studies, "packet radio" provides a low-cost alternative. This technique, pioneered by radio amateurs two decades ago, lets two computers share data without a phone or Internet link. The scheme is to break a data file into smaller chunks called packets, add routing information and send them over the airwaves. Although the transmitters used for this form of communication typically broadcast less than one watt of power, packet radio signals can be received almost anywhere in the U.S. through a nationwide network of amateur-operated stations (called repeaters) that automatically receive, boost and retransmit the information onward to more distant sites.

Packet radio has another advantage. Thanks to the visionary work of Bob Bruninga of the U.S. Naval Academy Ground Satellite Station in Annapolis, Md., generating a map on a personal computer that would locate all such drifters in real time would be a straightforward task. Bruninga created the Automated Position Reporting System (APRS), a software package that plots the data received on a digitized map and shows the precise location of the originating packet radio transmitter, be it on a plane in the air, on a ship at sea

DANIELS & DANIELS

"HERNDON ELEMENTARY," a research-grade drifter deployed by the Minerals Management Service, relayed measurements to government scientists and interested students for a month this past spring.

or even on a satellite orbiting in space. To find out more, consult www.aprs.org on the Web.

To use packet radio, the electronics assembly will have to include something called a terminal node controller (or, in the lingo of ham radio, a TNC), along with a GPS unit and an inexpensive transmitter. And at least one sophisticated TNC, the KPC-3 Plus from Kantronics (www.kantronics.com; 785-842-7745; $149.95), can simultaneously monitor two different analog sensors, convert the measurements into packets and transmit them. With minor modification, additional analog channels become available. (You'll have to be clever: Kantronics doesn't document or encourage making these changes to the circuitry.) So the drifter could, for example, measure water and air temperature, salinity, intensity of sunlight, and battery voltage.

For a transmitter, check out your local ham radio outlet—but be prepared to be dazzled. Most offer a smorgasbord of gear with a dizzying array of features. Remember, if everything goes well, in a month your transmitter will be lying in Davy Jones's locker, so stick with the simplest model that can do the job. A dedicated data transceiver (a radio that both receives and transmits) is probably the cheapest way to go. MFJ Enterprises in Starkville, Miss. (800-647-1800), for example, sells its model APRS MFJ-8621X2 for about $140. This unit transmits digital data on a standard APRS frequency. Or you may want to consider the Radio Shack HTX-202. It is more flexible and sells for only $199.95. Either unit should be detectable about 10 kilometers out to sea.

Remember, your drifter will need electricity for a month. But don't be seduced by the siren song of solar power: photovoltaic panels, rechargeable batteries and the electronics to connect them together will add complexity and at least $150 to the cost. Instead consider using ordinary alkaline D cells, which can last five times longer than rechargeables. By powering up the sensors and broadcasting its position only every half an hour or so, the drifter should be able to run for several weeks on six D batteries.

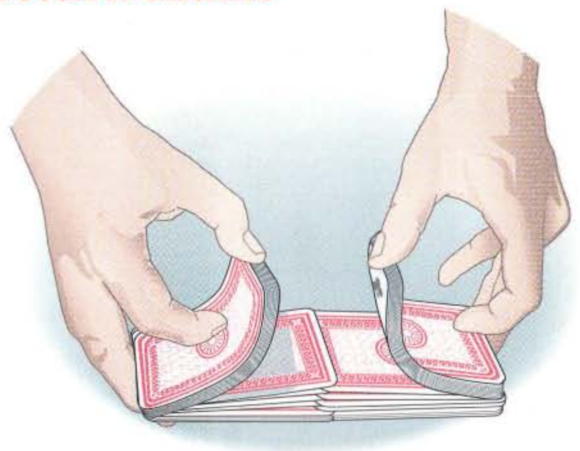
I expect that many of the clever people who read this column will provide some novel ideas for oceanographers to consider. The government agencies involved certainly want to enlist science enthusiasts: right now they offer a way for anyone to track some of their drifters over the Internet (check www.drifters.doe.gov on the Web). Perhaps, with the right design, amateur experimenters and high school students could assist the professionals by building coastal drifters and following their progress. Such activity would help forge a valuable alliance, enriching all parties while revealing some secrets of the sea. SA

I thank Ellen Prager, Bob Bruninga and Joe Valencic for their help and advice. For more about this and other projects, visit the forum hosted by the Society for Amateur Scientists at web2.thesphere.com/SAS/WebX.cgi on the World Wide Web. You may also write the society at 4735 Clairemont Square, Suite 179, San Diego, CA 92117, or call 619-239-8807.

MATHEMATICAL RECREATIONS

by Ian Stewart

Resurrection Shuffle

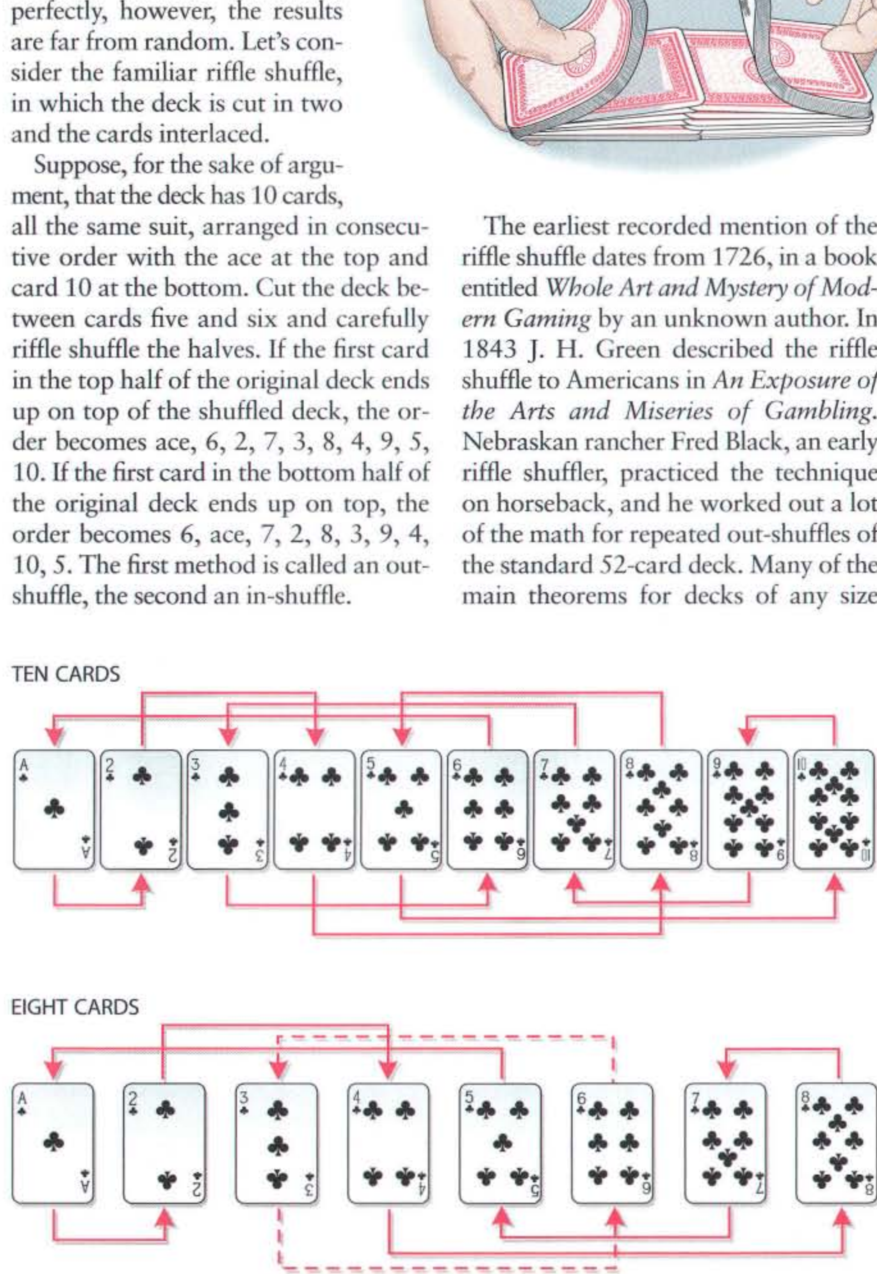


IAN WORPOLE

Most card games begin with someone shuffling the deck. The aim of shuffling, of course, is to randomize the order in which the cards appear. If the cards are shuffled too perfectly, however, the results are far from random. Let's consider the familiar riffle shuffle, in which the deck is cut in two and the cards interlaced.

Suppose, for the sake of argument, that the deck has 10 cards, all the same suit, arranged in consecutive order with the ace at the top and card 10 at the bottom. Cut the deck between cards five and six and carefully riffle shuffle the halves. If the first card in the top half of the original deck ends up on top of the shuffled deck, the order becomes ace, 6, 2, 7, 3, 8, 4, 9, 5, 10. If the first card in the bottom half of the original deck ends up on top, the order becomes 6, ace, 7, 2, 8, 3, 9, 4, 10, 5. The first method is called an out-shuffle, the second an in-shuffle.

The earliest recorded mention of the riffle shuffle dates from 1726, in a book entitled *Whole Art and Mystery of Modern Gaming* by an unknown author. In 1843 J. H. Green described the riffle shuffle to Americans in *An Exposure of the Arts and Miseries of Gambling*. Nebraskan rancher Fred Black, an early riffle shuffler, practiced the technique on horseback, and he worked out a lot of the math for repeated out-shuffles of the standard 52-card deck. Many of the main theorems for decks of any size were published in 1957 by British magician Alex Elmsley.



IAN WORPOLE

SHUFFLING A DECK of 10 cards reorders them in a 10-step cycle. Shuffling an eight-card deck reorders the cards in two cycles.

An out-shuffle can be viewed as an in-shuffle on a deck with two fewer cards. If we keep the ace at the top of our 10-card deck and card 10 at the bottom but perform an in-shuffle on cards two through nine, the order becomes ace, 6, 2, 7, 3, 8, 4, 9, 5, 10—the same order that resulted from an out-shuffle of all 10 cards. This connection allows us to consider just one of the two shuffles.

Suppose we take our 10-card deck and repeatedly subject it to an in-shuffle. What happens? Does the deck get more and more jumbled? At first the shuffling seems to randomize the deck—after three shuffles, the order becomes 7, 3, 10, 6, 2, 9, 5, ace, 8, 4. But after the fifth shuffle, the entire deck has exactly reversed its order! Clearly, five more riffle shuffles will "resurrect" the original order. We conclude that the in-shuffle, applied repeatedly to 10 cards, cycles through just 10 different orders. This is a tiny fraction of the 3,628,800 different ways of ordering 10 cards.

If you try the same procedure with any deck containing an even number of cards, you'll find that the deck always returns to its original order if the riffle shuffle is repeated enough times. Why is such repetition inevitable? The illustration at the left shows how each card in the 10-card deck moves during an in-shuffle. For example, the ace moves to the deuce's place, the deuce moves to card four's place and so on. Following the arrows, we see that the cards take one another's places in the following order: ace→2→4→8→5→10→9→7→3→6→ace. With each shuffle, the cards move one step further along this cycle. Because the cycle contains 10 steps, after 10 shuffles every card returns to its starting point.

The atypical feature of this deck is that there is just one such cycle. A more typical case is the eight-card deck, also shown at the left. Here there are two cycles: ace→2→4→8→7→5→ace, and

3→6→3. The first cycle repeats after six shuffles, the second after two. When the first cycle has reached its first repeat—after six shuffles—the second cycle has repeated for the third time.

However many cards there are, their movement through the deck can be broken down into a number of such cycles. Why? Start with any card and follow its progress. Because the deck is finite, eventually the card must reach a position it has previously occupied. From that stage on, it will repeat its previous moves. But can we be sure that when the card first repeats a previous position, it is repeating its initial position? The answer is yes, because any shuffle is reversible. To ensure that we can backtrack to the beginning, the first repeated position must be the initial position. For similar reasons, no card can jump from one cycle to another.

Once we know the cycles, there is a simple way to find out how many shuffles it takes to resurrect the original order of the deck. The length of each cycle is determined by its number of steps: if a cycle has *n* steps, each card in the cycle returns to its original position after *n* shuffles. If a deck has more than one cycle, we must determine the lowest common multiple of all the cycle lengths. For example, suppose a deck has two cycles, the first repeating after 10 steps and the second repeating after 14 steps. The cards in the first cycle return to their original positions at shuffles 10, 20, 30, 40, 50, 60, 70 and so on. The cards in the second cycle return at shuffles 14, 28, 42, 56, 70 and so on. The first number common to both sets, the lowest common multiple of 10 and 14, is 70. On the 70th shuffle all the cards will return to their original places.

Riffle shuffling will always resurrect the original order, no matter how big the deck. But is there a relation between the number of shuffles required and the number of cards in the deck? The fact that a 10-card deck repeats in 10 shuffles—equal to the size of the deck—is not typical. In fact, decks of 4, 6, 8, 10, 12, 14, 16, 18, 20, 22 and 24 cards, respectively, require 4, 3, 6, 10, 12, 4, 8, 18, 6, 11 and 20 in-shuffles to bring them back to their original order. There is a pattern here, but you have to be a number theorist to spot it!

Take the case of 10 cards. The table above shows successive powers of two

FOR TEN CARDS

EXPONENT (n)	1	2	3	4	5	6	7	8	9	10
POWER OF 2 (2^n)	2	4	8	16	32	64	128	256	512	1,024
REMAINDER OF $2^n/(10+1)$	2	4	8	5	10	9	7	3	6	1

FOR EIGHT CARDS

EXPONENT (n)	1	2	3	4	5	6
POWER OF 2 (2^n)	2	4	8	16	32	64
REMAINDER OF $2^n/(8+1)$	2	4	8	7	5	1

IAN WORPOLE

NUMBER OF IN-SHUFFLES
needed to resurrect the order of a deck can be found by dividing successive powers of two until the remainder is equal to one.

and the remainder obtained when each power is divided by 11 (the number of cards plus one). The remainder is one for the 10th power of two—and the number of in-shuffles needed to resurrect a 10-card deck is 10. For an eight-card deck, we look at the remainders obtained when each power of two is divided by nine. We get a remainder of one for the sixth power, and six is the number of in-shuffles needed to resurrect an eight-card deck.

This rule works in general. The number of steps needed to obtain a remainder of one is always equal to the number of in-shuffles needed to repeat the deck's order. And this number will always be less than or equal to the number of cards in the deck. Because an out-shuffle is the same as an in-shuffle on a deck with two fewer cards, a similar rule holds for the out-shuffle, but now you divide the powers of two by the number of cards minus one.

For a standard 52-card deck, 52 in-shuffles are needed to resurrect the original order, but only eight out-shuffles. Testing this result is tricky; even expert magicians find it difficult to perform an exact riffle shuffle, with the cards perfectly interlaced, and performing several in a row is nearly impossible. Working backwards is easier: simply deal the deck as if to two people, then stack their hands on top of each other. The reverse of an in-shuffle is called an in-sort, and the reverse of an out-shuffle is called an out-sort. The number of steps needed to resurrect the original order of the deck is the same whether you shuffle or sort. SA

FEEDBACK

In "What a Coincidence!" [June], I suggested that the appearance of the number 253 in two different problems about birthday coincidences was itself a coincidence. As several readers pointed out, it isn't! The first problem was, "How many people must be in a room to make it likely that two or more share a birthday?" The answer is 23 people, among whom there are 253 possible pairings. The second problem was, "How many people, in addition to yourself, must be in the room to make it likely that someone shares your birthday?" The answer is 253. The calculations I used to get those answers seemed unrelated, but Joseph Gerver of Rutgers University realized they are linked by an approximation.

The probability of a coincidence in the second problem is $1 - (364/365)^n$, where n is the number of people in the room besides yourself. The coincidence becomes likely when the probability exceeds $1/2$—which occurs when n reaches 253. The probability of a coincidence in the first problem is $1 - (364/365) \times (363/365) \times \ldots \times (365 - m + 1)/365$, where m is the total number of people in the room. When m is much less than 365, $(365 - m)/365$ is approximately equal to $(364/365)^m$. So the probability formula can be rewritten as $1 - (364/365)^{1+2+\ldots+m-1}$.

A comparison of the formulas shows that the two probabilities will be approximately equal when $1 + 2 + \ldots + m - 1$ is about equal to n. The number of pairings among m people is $m \times (m-1)/2$, which is equal to $1 + 2 + \ldots + m - 1$. So the number of pairings in the first problem must be approximately equal to the number of people in the second problem. —*I.S.*

REVIEWS AND COMMENTARIES

THREE MATERIAL WITNESSES

Review by James D. Livingston

The Substance of Civilization: Materials and Human History from the Stone Age to the Age of Silicon
BY STEPHEN L. SASS
Arcade Publishing, New York, 1998 ($24.95)

Stuff: The Materials the World Is Made Of
BY IVAN AMATO
HarperCollins, 1997 ($25); Avon Books (Bard), 1998 (paperbound, $12.50)

Understanding Materials Science: History, Properties, Applications
BY ROLF E. HUMMEL
Springer-Verlag, New York, 1998 ($59.95)


COPPER MINERALS


IRON DAGGER (IRAN, LURISTAN)


STONE AGE KNIFE, ARROW AND SPEAR TIPS


IRON AGE UTENSILS (CELTIC)

I teach in a department of materials science and engineering, a subject that was not taught, at least under that name, when I went to college. When in response to questions at cocktail parties I reply that I teach materials science, I am met with uncomprehending stares. Although my field is today one of the most important and dynamic areas of intellectual inquiry, it remains unknown and mysterious to most. These three recent books attempt to make the evolving field of materials science accessible to nontechnical readers.

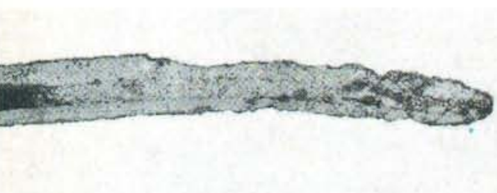
The early stages of civilization are commonly described in terms of materials—the Stone Age, the Chalcolithic (Copper-Stone) Age, the Bronze Age and the Iron Age. Early humans relied on natural materials, but we learned over the millennia that fire could convert clays into pottery, ores into metals and sands into glass, and we developed an ever wider assortment of materials with which to form our utensils, decorations, tools and weapons. Progress remained slow, however, until the 17th century, when Robert Boyle articulated the modern concept of chemical elements. By the 19th century most of the 100-odd elements of the periodic table (some odder than others) had been identified, providing a solid basis for the Industrial Revolution. And today's technology, from computers and lasers to jet engines and space probes, has been based on 20th-century advances in materials science.

Like musicians with their notes and painters with their colors, materials scientists experiment with new combinations and arrangements of the finite set of chemical elements to produce new and exciting things. According to a display at my department's office, the four major components of materials science and engineering are PRocessing, PRinciples, PRoperties and PRoducts. (The display was clearly created by an expert in PR.) In our field, the basic principles of chemistry and physics help us understand how the properties of engineering materials depend on processing methods used to manufacture them, with the ultimate goal of a product that can become part of a technological device.

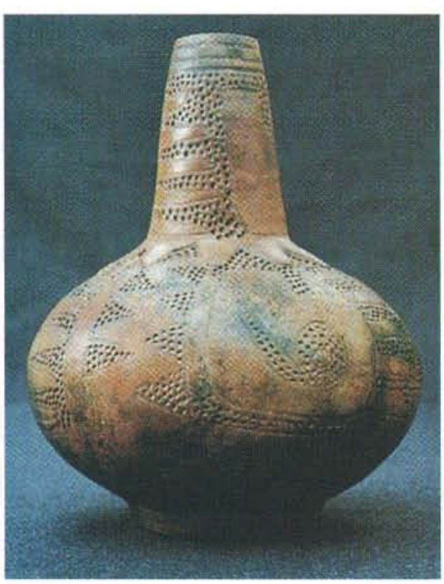
The intellectual foundation for the interdisciplinary field of materials science grew out of earlier work by metallurgists, ceramists, physicists and chemists to understand the properties of solids. In 1958 the department of metallurgy at Northwestern University formally changed its name to the department of materials science, and many other universities around the world soon followed suit. The Materials Research Society was founded in 1973 with 215 members and by 1996 had grown to a membership of over 12,000; the Federation of Materials Societies, an umbrella organization representing many large and small materials-related societies, claims about 750,000 members. Recognizing that all technological devices are limited by the properties of available materials, the Federal Coordinating Council on Science, Engineering and Technology has identified advanced materials as a national priority area for

ALL PHOTOGRAPHS FROM *Understanding Materials Science*

GOLD CRYSTALS



GLASS BOWL (ROMAN)



CLAY BOTTLE (NORTH AMERICA)

***EARLY MATERIALS AND OBJECTS* pictured here have given way to today's materials: steel, aluminum, plastics, composites.**

government support (along with global change, computing, biotechnology, math and science education, and advanced manufacturing). Materials science has clearly arrived.

In his introduction to *The Substance of Civilization,* Stephen L. Sass writes that he seeks to address the question "How did materials shape our culture?" He accompanies his description of the development of early civilizations with numerous materials-related quotations from the Old Testament, the *Odyssey* and other ancient sources. For example, he cites Deuteronomy 8, in which Moses describes the Promised Land as "a land whose stones are iron, and out of whose hills you can dig copper," adding that "Moses knew that iron and copper would be as enticing to his followers as milk and honey." Later in the book, he discusses composites (materials, such as modern fiberglass, produced from two dissimilar materials to achieve properties superior to either component) with a quote from Exodus 5 describing the use of straw to improve the strength of brick. Sass, who teaches materials science at Cornell University, cannot resist diverting from the history of civilization occasionally to teach a few of the relevant basics of his field. But he does this gently, without the use of equations, and only the most extreme sciencephobes will find this troublesome.

In the last third of his book, Sass focuses on modern materials, including steel, aluminum, plastics, composites and man-made diamonds, ending with a chapter on "The Age of Silicon," the semiconducting material that is the basis of the integrated circuits that run our computers and much of modern technology. The prime strength of his book, however, lies in the earlier chapters, which point out the close relationship of materials developments to advances in early civilizations. The readability of this book is enhanced with occasional personal references: the author opens his chapter on the history of glass by recalling the day his son broke a window with a misdirected tennis ball. Although the author and his publisher committed the unforgivable sin of omitting an index, *The Substance of Civilization* indeed contains much of substance and is a good starting place to develop an appreciation for the history and nature of materials science.

Stuff was written by Ivan Amato, an experienced science writer who in his acknowledgments reports being inspired by the late Cyril Stanley Smith of the Massachusetts Institute of Technology, "whose poetic and philosophic approach to the world of materials has changed me forever." Amato has become an enthusiastic proponent of the importance of materials science, which he calls a "superdiscipline." Although he also records some of the early history of materials, his focus is on the frontiers of modern materials science. A colorful writer, Amato includes in his description of today's "materials Serengeti" in-depth discussions of synthetic diamond, biomimetic materials (man-made materials modeled after natural biological materials) and smart materials (ones that sense pressures, temperatures or other conditions and alter themselves to deal with them). He discusses modern processing ("spray painting with atoms") and modern analytical microscopy, which has advanced so much in resolution that individual atoms can now be seen, as he demonstrates with a few well-chosen microphotographs.

The one-word career tip—"Plastics"—given to Dustin Hoffman in the 1967 film The Graduate *was actually very good advice.*

In two chapters based on extensive interviews with practicing materials scientists, Amato describes modern processing and analysis of sophisticated semiconductor "quantum cascade lasers" and of steel, the seemingly mundane but exquisitely complex material that shapes much of our world. He ends, like Sass, with a speculative glance into the advanced materials of the future, which he calls "the stuff that dreams are made of." Amato explains some of the science and technology of materials in more depth than Sass does, but his lively writing style keeps his descriptions of even the subtleties of semiconductor physics and steel metallurgy enjoyable.

In *A Brief History of Time,* Stephen W. Hawking reported being told that each equation he included in his book would halve the sales, a warning apparently noted by both Sass and Amato,

who avoid displaying any equations (although one or two are discreetly described in words). Nevertheless, for those unthreatened by mathematics, an equation can be worth a thousand words. Such readers could appreciate both parts of Rolf E. Hummel's *Understanding Materials Science,* which is actually two books in one: a nonmathematical description of the history of materials and an introductory textbook complete with equations, graphs and instructive diagrams.

Hummel follows a description of the Stone and Chalcolithic ages with a basic scientific discussion of the strength of materials; he ends the Bronze Age with a discussion of strengthening mechanisms in alloys; and he closes the Iron Age with a look at the metallurgy of iron and steel. He considers the present "the Age of Electronic Materials" and, after a historical interlude, discusses the electrical, magnetic, optical and thermal properties of solids. Noting that no Ceramics Age has been identified because *every* age of civilization has utilized ceramics, Hummel then tackles the history and science of pottery, glass and concrete and follows that with a chapter on fibers and plastics. (Both Sass and Amato note that the one-word career tip—"Plastics"—given to Dustin Hoffman in the 1967 film *The Graduate* was actually very good advice.)

In the center of *Understanding Materials Science* is a series of fascinating color plates showing early objects of pottery, metal and glass from various museums. Hummel teaches materials science at the University of Florida, where this book will presumably be used as a text. In his preface, however, he expresses the hope that his book will also attract a sizable readership from the humanities. Such readers would enjoy reading the historical sections, studying the color plates and at least skimming the more technical chapters to get a quick overview of the basic science of materials. Unfortunately, the book carries the price of a textbook rather than that of a trade book, which will limit its appeal to a general audience.

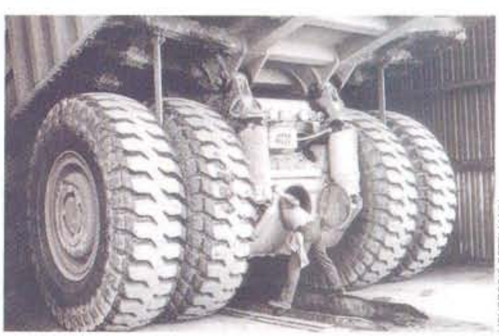
As Madonna sang in what became her trademark song, "We are living in a material world (and I am a material girl)." Yet few of us in this material world appreciate the impact of improved materials on modern technology or give much thought to the myriad of materials we use every day.

And the relative newness of materials science and engineering as an organized field of study brings with it a problem for teachers like me who specialize in undergraduate education. Because the subject is unfamiliar to most students entering college (and to their parents), convincing undergraduate students to major in materials science is a challenge. (In contrast, attracting graduate students is easy; by the time they have finished their undergraduate studies, most science and engineering students have discovered the importance of materials science.) The three books reviewed here, by demonstrating to a general readership the overarching importance of materials to the history of civilization and of modern materials science to future technological progress, may lead to more freshmen arriving on campus with an interest in this engrossing topic.

JAMES D. LIVINGSTON is a senior lecturer in the department of materials science and engineering at M.I.T. He is the author of Driving Force: The Natural Magic of Magnets *(1996).*

THE EDITORS RECOMMEND

HEAVY EQUIPMENT. Erik Bruun and Buzzy Keith. Black Dog & Leventhal Publishers, New York, 1997 ($25).



UPI/CORBIS-BETTMANN

A big book ($13\frac{3}{4}$ inches square) describes big machines, providing eye-popping pictures of the monsters at work and good explanations of how they lift, dig, crush, mine, push earth and carry huge loads. One learns that the Komatsu 575A-2, pictured on page 21, is the world's largest bulldozer, capable of moving 90 cubic yards of dirt with a single shove. And that a massive, electric mining shovel needs a 7,200-volt power supply and, working 450 hours a month, will consume 430,000 kilowatts of electricity at a cost of $30,000. David J. Bowler, who operates several kinds of heavy equipment, gives the reader a sense of the skill and care required: "The day I think I know everything about running a crane is the day I'm going to get into trouble. Every job is different. You've got to know your load and you've got to know the machine's limit."

THE CRUCIBLE OF CREATION: THE BURGESS SHALE AND THE RISE OF ANIMALS. Simon Conway Morris. Oxford University Press, New York, 1998 ($30).

The Burgess Shale, a thin outcrop of rock in the Canadian Rockies, contains a rich store of extraordinarily well preserved fossils of creatures that lived in the Middle Cambrian period, 500 million years ago. The fossils have provided a vital key to understanding the early evolution of animal life. Conway Morris, professor of evolutionary paleobiology at the University of Cambridge, has explored the shale since 1972. He describes the scene and the fossils vividly, using the device of a time machine that takes a group of scientists back to the Middle Cambrian and disgorges a small submersible wherewith they venture into the sea to view the creatures as they looked and acted in life. But he has a further purpose, which is to dispute the interpretation that some other scholars—notably Stephen Jay Gould of Harvard University—have put on the evolutionary significance of the Burgess Shale animals. Gould, he says, argues that if the tape of life were rerun from Cambrian time, we would end up with an entirely different world, which would include among its various features the absence of human beings. "On the contrary," Conway Morris writes, "I believe it is necessary to argue that within certain limits the outcome of evolutionary processes might be rather predictable."

THE GIFT OF TIME: THE CASE FOR ABOLISHING NUCLEAR WEAPONS NOW. Jonathan Schell. Metropolitan Books, New York, 1998 ($25).

The gift of time—time propitious for reaching an international agreement to abolish nuclear weapons—comes from the end of the cold war. Schell presents his argument in his own words and in inter-

views he had with several prominent men who were once engaged in framing nuclear policy and planning nuclear war—among them Robert S. McNamara and Mikhail S. Gorbachev. He also states clearly the counterargument that nuclear weapons cannot be disinvented and that policing the agreement would require that we have "a verification regime of extraordinary rigor and intrusiveness."

Schell frames the debate in these terms: "Was the nuclear buildup a story of wise management of a terrible dilemma that had no other solution? And was it, further, a fortuitous training session, in which the world was introduced to the previously unsuspected virtues of nuclear arms? Or was it a tale of the reckless endangerment of mankind? If the first is true, then nuclear weapons are a marvelous gift of proven worth to the world. If the latter is true, then abolishing nuclear weapons is the unfinished business of the end of the Cold War. In the first case, we *cannot* do without them. In the second, we *must* get rid of them." Schell argues the second case eloquently and asserts that "it is possible, if only we can muster the political will, to free the world of nuclear danger."

NIGHT HAS A THOUSAND EYES: A NAKED-EYE GUIDE TO THE SKY, ITS SCIENCE, AND LORE. Arthur Upgren. Plenum Press, New York, 1998 ($27.95).

American artist James McNeill Whistler was invited one sparkling night to step outside and view the panoply of stars. No, he said, "There are far too many of them and they are so very poorly arranged." Upgren, an astronomer at Yale and Wesleyan universities, sets out to make star viewing easier for everyone else who feels as Whistler did and for those who admire the beauty of the night sky but know little about what they see there. He does it clearly and smoothly, describing first the principal stars and constellations as they appear in each season of the year and then the sun, moon and planets.

Along the way he supplies an abundance of related facts (the Big Dipper is properly called an asterism, "a name for an easily noticeable group of stars that does not make up a full constellation"; the stars Mizar and Alcor in the Dipper have "long been recognized as a test for good eyesight in many cultures"), discusses Stonehenge and the mysterious Old Stone Mill in Newport, R.I., takes a dig at astrology ("Your horoscope is two thousand years out of date" because the astrological arrangement of the signs of the zodiac was keyed to the celestial alignment of Roman times), and makes a plea for reducing the glare of urban lights that impedes star viewing by astronomers as well as by Whistler types.

MOOD GENES: HUNTING FOR ORIGINS OF MANIA AND DEPRESSION. Samuel H. Barondes. W. H. Freeman and Company, New York, 1998 ($24.95).

The search for genes that cause or contribute to disease has had many successes in recent years, and the successes come at an ever faster pace. Barondes, a psychiatrist who has treated many patients with serious mood disorders, examines here the prospects for finding genes that cause or contribute to such afflictions—in particular manic-depressive illness. "Characterized by episodic and disruptive mood fluctuations," he says, "this illness is especially important because it affects so many people—about one in a hundred of us in its most flagrant form, and possibly several times as many of us in milder versions."

His tale focuses on two families in which the disease has turned up repeatedly. Having introduced the first of the families, he provides what amounts to a basic course on genes, presented so clearly as to make the reader feel almost like an expert on the subject. Barondes is optimistic that mood genes will be found, thus introducing the possibility of "tailor-made treatments for different forms" of mood disorders and of "the development of new therapeutic drugs that alleviate symptoms and prevent attacks."

APES, LANGUAGE AND THE HUMAN MIND. Sue Savage-Rumbaugh, Stuart G. Shanker and Talbot J. Taylor. Oxford University Press, New York, 1998 ($29.95).

Kanzi, a male bonobo (an ape sometimes called a pygmy chimpanzee), has been under the care of language researcher Savage-Rumbaugh since infancy. Over a period of 18 years, he has learned to communicate his wants and to respond to spoken English by means of pictorial symbols called lexigrams. His communicative capability is about equal to that of a two-and-a-half-year-old human child. The first third of the book presents Savage-Rumbaugh's clear and entertaining account of Kanzi's upbringing. The remainder, largely written by the other two authors, is an argument in academic prose addressed primarily to critics who "insist that no ape has ever developed truly linguistic skills." The authors declare their "shared belief that the Kanzi research presents a serious and effective challenge not only to scientific thinking about the cognitive and communicational capacities of nonhuman primates, but also to received knowledge concerning the possession of those capacities by humans."

ANNA CLOPET

HISTORY OF THE HOUR: CLOCKS AND MODERN TEMPORAL ORDERS. Gerhard Dohrn-van Rossum. University of Chicago Press, Chicago, 1996 ($29.95).

Dohrn-van Rossum, who teaches medieval and early modern history at the University of Bielefeld in Germany, has researched his subject so profoundly that he can append to this book 81 pages of notes—mostly bibliographical—running to 948 entries. He treats the subject profoundly, too, dealing not only with the history of timekeeping devices from the sundial to the cesium clock but also with changes in the human conception of time from the cyclical order of "Church's time" to the linear order of "merchant's time." The prose sometimes plods, and key points are not always crisply stated, but the story of timekeeping is here in wonderful depth.

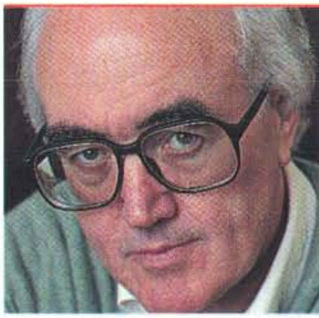
BIBLIOTHÈQUE NATIONALE, PARIS

ISAAC NEWTON: THE LAST SORCERER. Michael White. Addison-Wesley, Reading, Mass., 1997 ($27).

Most scholars have known that Newton had rather unorthodox intellectual pursuits that seemed to run counter to his work in empirical science. Yet biographers refused to reconcile the paradox. Until Michael White. White elucidates both sides of the debate about whether Newton's alchemical studies cross-fertilized his scientific research, but his conclusion is that Newton's search for the philosophers' stone and for scientific fundamentals such as gravity are "inextricably linked."

Moving beyond the debate, White draws on numerous letters and documents to describe many of Newton's relationships—good and bad—with people such as Edmund Halley, Robert Hooke and Nicholas Fatio de Duillier, with whom it is believed he had a sexual liaison. "Newton was a man driven by childhood pain, but his subconscious method of dealing with this angst was to accumulate power and to gain dominance over others." These characteristics are even more obvious after Newton's nervous breakdown in 1693, when he discarded academic pursuits for more heavy-handed work as a "private investigator and prosecutor" who was feared by many.

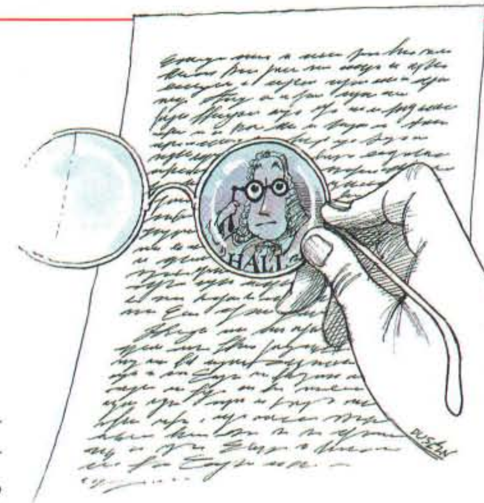
COMMENTARY



CONNECTIONS

by James Burke

Oops



DUSAN PETRICIC

From time to time it gives me great pleasure to come across, and publicize, the name of somebody who never got the credit. In this case, peering through my optically pure reading glasses in the British Library, I spy the name of a person who, in the privacy of his later years, must have said at least once, "Oops." Chap called Chester Moor Hall. Who? Okay, try this one: John Dollond. Right! The guy who invented the optically pure achromatic lens in 1757? Nope. Hall did it, years before.

A nice guy, Hall—and like many such, he finishes last. Around 1729 he becomes convinced that it is possible to make a lens that will not give you all kinds of blurred, colored images. The bane of astronomers up to then, and the reason they are seeing things like "planets with ears" (Saturn). So by 1733 Hall sticks lenses of differing densities (flint and crown glass) together, and the different dispersions kind of cancel one another out. Bingo. Achromatic, no-color fringes, total clarity. He makes a couple of telescopes for friends, sticks the whole experiment in the closet and goes back to being a landowning magistrate in Essex. Never does anything about it. Even when he hears some other guy (name of Champness) fussing about having invented it before Dollond. For the news of Chester Moor Hall's work to get mentioned in a paper to the Royal Society (where it's then filed and forgotten) takes another 100 years. And then another 165 to get to me. Forefront of historical research I'm not. But you already knew that.

Back to John Dollond. He gets out of silk weaving (the d'Hollande family had originally been Dutch textile manufacturers) into optics. He and his son set up shop in Piccadilly, London, and become instrument makers to the gentry. Dollond's fortune is made. On his death, he leaves his patents for many inventions, including the achromatic lens, split among his heirs, one of whom is his apprentice and the husband (no fool) of his daughter, Sarah: Jesse Ramsden.

For Ramsden, this is a marriage made in heaven. He has his hands on the lens, plus he's a real whiz with metal, particularly when it comes to making marks on it. Degrees and minutes of arc, that kind of stuff. Which he proceeds to do on the 1,000 sextants he then produces for the British navy and various explorers. The thing about Ramsden's marks is that they're extremely accurate. And the more accurate the marks on your instrument, the more accurate your measure of whatever it is you're measuring. In the case of navigators, using a Ramsden-marked sextant means you're more likely to avoid the rocks.

As long as you know the rocks are there in the first place. Not the kind of knowledge that is widespread, back in

He tried making artificial diamonds by heating iron and carbonized sugar.

the days when the Industrial Revolution is beginning to require the import of thousands of tons of near-freebie raw materials from new colonies and the export of finished goods to the same. All the Brits have to do to facilitate this sweetheart import-export deal is pass laws to oblige the hapless colonists to sell to, and buy from, only the mother country. And do it all in only British ships. One reason the U.S. is now the U.S. Still, the scam works long enough elsewhere (India, Africa) to fund most of the culture and stately homes the modern tourist now enjoys in Britain. All these profits are why lots of heavily laden, low-in-the-water ships are racing back and forth, all too often getting lower in the water than you might wish (after hitting the aforementioned rocks). Hence the great lighthouse-building mania of the period. Now, putting up lighthouses has been no big deal since Pharos. The problem for the late 18th century isn't the "house" bit. It's the "light" bit. As in candles that nobody can see until it's too late.

Because this is a bottom-line matter, things get done about it. Thanks to Jesse Ramsden's accurate marks. In this case, on a giant, four-foot theodolite that makes possible the Ordnance Survey of Ireland. Once, that is, the surveying teams can see what it is they're pointing the theodolite at. Not easy early on, in the murk of Ulster. Until 1825, when enters a young army type, Thomas Drummond, who comes up with one of those gizmos that go into the language. A jet of hydrogen and oxygen burns and is directed at a small ball of lime. As the flame becomes incandescent, it is reflected by a parabolic mirror. And voilà, limelight. Great for surveyors, actors and mariners heading for the rocks. Unless the lighthouses run out of gas. They do.

By 1849 Belgian professor Floris Nollet can make all the gas you might ever want, with electrolysis. Place positive and negative electrodes in water, and the current between them dissociates hydrogen and oxygen from the water. Problem solved. Unless the lighthouses run out of electricity. They do. Advances are made (remember, this is a bottom-line matter), and by 1871 a co-worker of Nollet, Zénobe Théophile Gramme, offers the dynamo. Wire coils spin in a

magnetic field and generate electricity. Enough to power an arc light. Two carbon rods almost touch. As electricity coming through the rods jumps the gap between them, sparks fly, and the carbon points go incandescent. Makes limelight look dark. And paves the way for the electric arc furnace, in which you can use two carbon electrodes to create humongously hot heat.

Enough for a Frenchman in 1892 to believe he can make artificial diamonds by heating iron and carbonized sugar at temperatures so great the carbon will dissolve into the iron. He will then rapidly cool the iron in water, so it solidifies with enormous pressure, causing microscopically small carbon particles to appear. The furnace inventor and would-be diamond merchant Henri Moissan thinks what he comes up with are the aforementioned synthetic stones. As it turns out, they aren't.

But nobody cares, because the arc furnace has dazzled the scientific community far more than Moissan's little carbon chips have. In 1894 Moissan puts a mix of lime and carbon in the furnace, heats it to 2,000 degrees Celsius and gets some pretty dull stuff until he brings it into contact with water, when it gives off a gas. Acetylene. Which is also pretty dull, until he sets light to it. Makes arc light look dark. By 1899 there is a proliferation of acetylene plants, most of them in places such as Niagara Falls, the Pyrenees, Norway and Switzerland, where plenty of falling water generates the electricity required for the furnaces. Then comes Thomas Edison. When he does what he does, the bottom drops out of the acetylene-light market.

By 1912 lonely acetylene freaks are noodling around looking for anything they might be able to do with the stuff. In Frankfurt a chemist at Greisheim Electron, on the alert for some kind of doping material with which to weatherproof the fabric on aircraft wings, tries a mix of acetylene, hydrogen chloride and mercury. No good. So he puts the goop aside on a sunny windowsill and notices that it forms a milky sludge and then goes solid. He files a patent and forgets it. And in 1925 the patent lapses. Which is why I end, as I began, with another guy whose favorite word must have been "oops." He was named Fritz Klatte. The sludge he ignored turned out to be polyvinyl chloride, PVC. SA

Wonders, continued from page 96

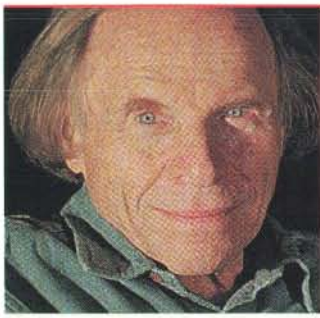
What would old Mersenne make of the present largest prime, one of his formula, reported early in 1998? It has nearly a million decimal digits. The printed number alone would about fill a good-quality floppy disk, a bookful. A big team (formed in 1996) is finding a few such primes every year. The 4,200 amateurs and dozens of professionals engaged in the Great Internet Mersenne Prime Search share special software and accept assigned ranges of search, where they seek out new primes in unattended runs. Their pooled results amount to a vast volunteer supercomputer, thousands of Pentiums hard at work all night on a single task. Another conjectural formula for huge primes predicts one with about 10^{37} digits, alas!

Since about 1900 we have known not all the primes one by one as we count up but the measure of their prevalence—the larger the number, the greater the accuracy. That 1998 record prime is one among a few million composite numbers nearby, a rare and delicate flower blooming in that prime-number desert.

Or maybe not. In the December 1958 issue of *Scientific American*, David Hawkins, a true philosopher, educator, mathematical amateur and my old friend, exhibited his iconoclastic variant on venerable Eratosthenes: a random sieve. Don't get out your calculator to find factors of any number you test; instead try random choice. First sieve out numbers at a 50–50 rate, simply tossing a fair coin. Then test your luck again, but now employing odds for survival set by the luck of the draw. In other words, if the number 4 is the next lucky survivor, set the odds of survival at one out of four.

Keep it up as far as you wish, shaping both the odds and actual survival by chance alone. What survives shows no numerical strictness at all; about half are even numbers, and no two big siftings repeat. Yet their dilution matches the certified primes. It is not the miraculous avoidance of factors all the way back to the number 2 that seems to fix the population of primes; it is their very sparseness of survival, provably mimicked by blind chance.

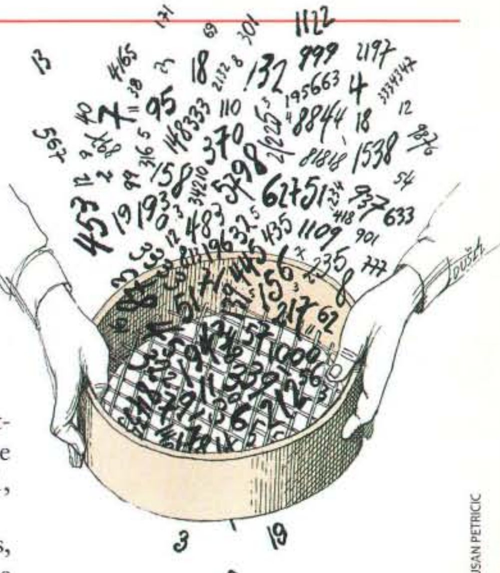
For hot news on primes, a rich book compendium and much lore, consult www.utm.edu/research/primes on the World Wide Web. SA



COMMENTARY

WONDERS
by Philip Morrison

Numbers: Prime or Choice?



DUSAN PETRICIC

The simplest infinity is the one already recognized by many a curious young girl or boy counting 1, 2, 3, 4…. My friend Carl Sagan recalled a Sunday long ago when at age six he had delighted in newfound counting skills. His father explained to him that there need be no end to the count. Why stop? Just add one more. Gung ho, the boy set out to write integers that very evening all the way up to a big—if finite—1,000! His father generously wrote on for the boy during the time lost to Carl's customary evening bath. Together father and son completed their kilocount by bedtime, an experience Carl never forgot.

Positive integers are an infinite family whose kinship is wonderfully intricate. Most integers can be generated by multiplying together enough smaller integers, called factors. Thus, the number 18 = 2 × 3 × 3. (But try 17!) Here a chemical metaphor arises: 18 is a molecule among numbers, a composite of three number-atoms, whereas 17 suggests a pure element, all one kind of atom. It is called a prime number and has as factors only itself and unity.

Primes are the atoms of multiplicative arithmetic. There are plenty of them, more than there are all material atoms. Euclid of Alexandria himself published a proof in the earlier part of the third century B.C. that the primes are unending. Another celebrated savant of Alexandrian times, Eratosthenes, developed a systematic scheme called a sieve to sift out the composites, leaving every prime standing. Try a tiny example. Write the numbers 1 through 20 in a row across your papyrus. Ignore that inert 1. Leave 2 alone, because it is a prime. Now strike out every second number to come, for they all have 2 as a factor. There go the nine even numbers, 4 through 20, but 3 survives. Repeat the process for 3, and this time 9 and 15 sift out. A prime itself, 5 has nothing left to do. Your sieve retains the pure primes: 2, 3, 5, 7, 11, 13, 17 and 19.

Had you tested the first 100 numbers, just a quarter would remain, pure prime content from 2 up to 97. Note that you need sift only with primes no larger than the square root of the number being tested. Any factor larger than the square root must have had one or more smaller factors to complete the test product. Primes thin out: there are eight up through 20 but only three between 80 and 100. Even more convincing is the sifting process itself: as you go on, each surviving number has resisted more and more tests. This progression resembles a sieve with many screens, each succeeding layer pierced by fewer and fewer openings.

Primes thin out: eight up through 20 but only three between 80 and 100.

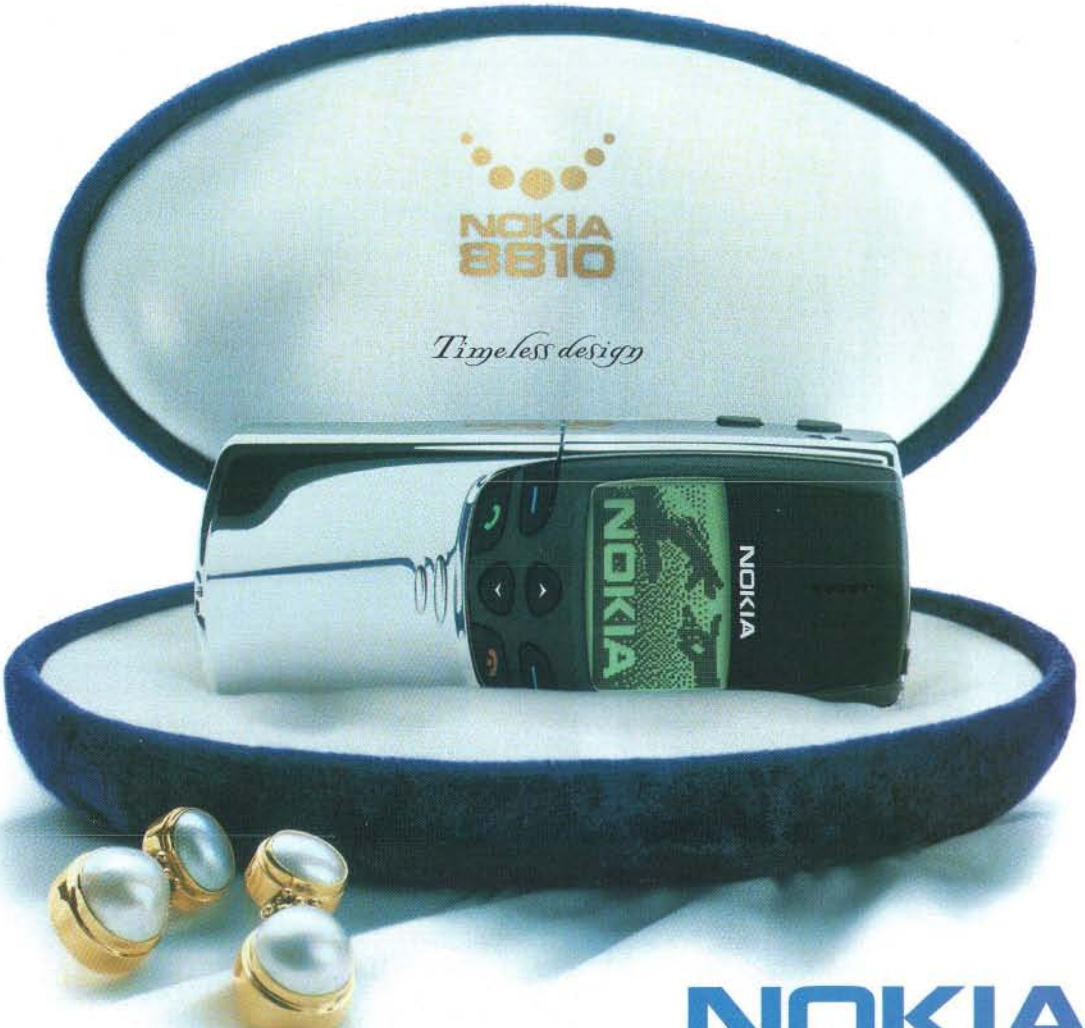
The unendingly surprising results of the theory of numbers continue to fascinate. Most big primes are found only by special theorems; the sieve performs a tediously clean sweep. A wealth of simply stated integer relations have been proved and many more conjectured without proof. For instance, Christian Goldbach asserted famously in 1742 that all even numbers greater than 2 are the sum of two primes, as in 96 = 89 + 7. No proof is known, yet no exception has been found either.

Over the past decade or so, prime numbers have become highly important to the growing technology of encryption for a wired world. Some feeling for the technical architecture of the best-known scheme of today, the public-key system, can be gained even from our mere sketch. A message can be regarded as a long number, a string of digits. To convey it in encrypted form is simple, given enough computing power: raise it to a preset exponent, divide the result by the long number that is the public key, then send the remainder. The encrypting composite number is all that need be made public, along with some digits that determine the processing. To reverse the encryption, a powerful theorem of the great Leonhard Euler allows calculation of the message number using the prime factors of the key but cannot work given only the composite key itself.

The factoring of seriously large numbers remains hopelessly slow. The present encryption scheme therefore seems foolproof, until the art of factoring 100-digit numbers matures.

Such widespread usefulness for large primes is brand-new; it rests on the extraordinary power of computing. Father Marin Mersenne, a 17th-century master, wrote that "all time would not suffice" to test a given number of 15 or 20 digits for primality. He then laid out a helpful formula for numbers he thought most likely to be primes: raise 2 to any prime power, then subtract 1 from that. The first few prime numbers used as exponents all give Mersenne primes, but the number formed with 211 is not prime. (No formula at all is yet known that surely leads to new primes.) Mersenne's conjectured primes are especially easy to test by computer; theorists have learned much about his formula, steadily finding clues to much larger primes.

Continued on page 95


NOKIA
8810
Timeless design
NOKIA
NOKIA